



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 133468

TO: Phuong Bui  
Location: REM/2A15/2C18  
Art Unit: 1638  
Monday, September 27, 2004

Case Serial Number: 10/624061

From: Deirdre Arnold  
Location: Biotech-Chem Library  
REM 1A64  
Phone: 571-272-2532

Deirdre.Arnold@uspto.gov

### Search Notes

## RUSH

Results are for SN 10/624,061.

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:32:10 ; Search time 68 Seconds  
(without alignments)

1211.033 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339  
Sequence: 1 WAKAEGSPDSGNVYTLASV.....NWVPLTKVEQVYLLRLRD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	818	61.1	265	10 Q9S7H4	Q9S7H4 arabidopsis
2	691	51.6	255	10 Q9STB2	Q9STB2 lycopersico
3	685.5	51.2	294	10 Q93160	Q93160 oryza sativ
4	689.5	50.0	316	10 Q9C544	Q9C544 arabidopsis
5	654.5	48.9	316	10 Q9XF60	Q9XF60 arabidopsis
6	432	32.3	295	3 Q96V28	Q96V28 cryptococcu
7	423	31.6	280	3 Q9P4D8	Q9P4D8 pichia angu
8	414	30.9	267	3 Q9Y7B2	Q9Y7B2 emericella
9	329	24.6	131	10 Q65861	Q65861 prunus aviu
10	283	21.1	82	10 Q22410	Q22410 petroselinu
11	188	14.0	70	10 Q22409	Q22409 petroselinu
12	110	8.2	4057	5 Q81J57	Q81J57 plasmodium
13	107	8.0	355	2 Q9AHC6	Q9AHC6 streptococc
14	105	7.8	360	16 Q975J4	Q975J4 streptococc
15	98.5	7.4	1259	17 Q8PU05	Q8PU05 methanosarc
16	98.5	7.4	1506	16 Q8EX35	Q8EX35 shewanella

#### ALIGNMENTS

#### RESULT 1

Q9S7H4 PRELIMINARY; PRT; 265 AA.

AC Q9S7H4;  
DT 01-MAY-2000 (TREMELrel. 13, Created)  
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Chorismate mutase (EC 5.4.99.5) (Putative chorismate mutase CM2).  
GN CM2 OR T3ON20.140 OR ATSG10870.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kuhn R., Vogt E., Schmid J., Amrhein N., Schaller A.;  
RT "Expression analysis of Arabidopsis thaliana genes for plastidic (CM1)  
and cytosolic (CM2) chorismate mutases";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97111372; PubMed=8953244;  
RA Amrhein N., Ehrler T.T., Eppe P., Felix G., Raesecke H.R.,  
RT "Cytosolic and plastidic chorismate mutase isozymes from Arabidopsis  
thaliana: molecular characterization and enzymatic properties.";  
RL Plant J. 10:815-821(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,  
Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Arabidopsis sequencing project;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.

17 96.5 7.2 668 10 Q8LPF1  
18 96 7.2 459 16 Q8CNG6  
19 95.5 7.1 565 16 Q8DLU6  
20 95.5 7.1 587 16 Q99RK3  
21 95.5 7.1 587 16 Q8NV02  
22 95.5 7.1 1556 3 Q08554  
23 95 7.1 281 16 Q8A4M0  
24 95 7.1 746 11 Q8C9T7  
25 94.5 7.1 761 16 Q89AS4  
26 94 7.0 1561 4 Q9ULG1  
27 93.5 7.0 1134 17 Q9UY58  
28 93 6.9 422 16 Q8XNP6  
29 93 6.9 958 10 Q8GUL2  
30 92 6.9 296 3 Q9UVD5  
31 92 6.9 371 16 Q9ZDP9  
32 92 6.9 812 16 Q8G017  
33 92 6.9 823 16 Q8YHC6  
34 92 6.9 1844 5 Q8IE31  
35 91.5 6.8 581 9 Q8SBQ2  
36 90.5 6.8 333 17 Q8Q3T5  
37 90.5 6.8 792 13 Q90ZB6  
38 90 6.7 274 2 Q47742  
39 90 6.7 660 5 Q8IIX6  
40 90 6.7 1015 5 Q94705  
41 89.5 6.7 251 16 Q51460  
42 89.5 6.7 405 16 Q8DTJ3  
43 89.5 6.7 534 2 Q9ZFS4  
44 89.5 6.7 1176 16 Q8RDW3  
45 89.5 6.7 1187 10 Q9ZV43

Q9LPF1 arabidopsis  
Q8CNG6 staphylococ  
Q8DLU6 synchococ  
Q99RK3 staphylococ  
Q8NV02 staphylococ  
Q08554 saccharomyc  
Q8A4M0 bacteroides  
Q8C9T7 mus musculi  
Q89AS4 buchnera ap  
Q9ULG1 homo sapien  
Q9UY58 pyrococcus  
Q8XNP6 clostridium  
Q8GUL2 arabidopsis  
Q9UVD5 nilaparvata  
Q9ZDP9 rickettsia  
Q8G017 brucella su  
Q8YHC6 brucella me  
Q8IE31 plasmodium  
Q8SBQ2 bacterioph  
Q8Q3T5 pyrococcus  
Q90ZB6 xenopus lae  
Q47742 enterococcu  
Q8IIX6 plasmodium  
Q94705 physarum po  
Q51460 borrelia bu  
Q8DTJ3 streptococ  
Q9ZFS4 bacteroides  
Q8RDW3 fusobacteri  
Q9ZV43 arabidopsis

#1, 6107.







QY 252 QVEYLLRRLD 261





DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR000787; Cor/por Metransf.  
DR InterPro; IPR007345; PS\_pyruv trans.  
DR InterPro; IPR001865; Ribosomal S2.  
DR Pfam; PF04230; PS\_pyruv trans; 1.  
DR PROSITE; PS00962; RIBOSOMAL\_S2\_1.  
SQ SEQUENCE 355 AA; 40777 MW; 60BDC233A233D6E7 CRC64;  
  
Query Match 8.0%; Score 107; DB 2; Length 355;  
Best Local Similarity 23.8%; Pred. No. 1;  
Matches 53; Conservative 36; Mismatches 76; Indels 58; Gaps 10;  
  
QY 31 IIVGLIERAKFPNSHT-----YDEKIAIQGFCGSLVEFVVVNTAIOAKAGRYKXP 83  
DB 21 IVDLLESKGYEASFPNPSDFQMYLKEYRQKQ-----SFTKQADAILYIPGGYFGE 73  
QY 84 EENAFFPENLPPSIVPSYFQFLHPGAASININKSIWKYFKELPLLATSGDDGNVY 143  
DB 74 GHNARFRDNL-----IQKRFPLG-----IWASYFKKPIGVLGAGPN----- 113  
QY 144 TAANDLSLQISRIHYGKFAEVRKFD-----AP--QDYEPILIRAKDKEGLM 190  
DB 114 ---NDSLWNYGKRIINHAQFIT--VRDRESFDSLKHLSPSPVHETFDLIISKLREKT 169  
QY 191 KLTTFTSVEETVRKRV-----EKKAV-VFGQEVNLSDDNDN 226  
DB 170 EQLCOLKREAKDKKIILVYHNHKKALEKFAESISLFLNPN 212  
  
RESULT 14  
Q975J4 ID Q975J4 PRELIMINARY; PRT; 360 AA.  
AC Q975J4;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein SP0355.  
GN SP0355.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC BAA-334 / TIGR4;  
RX MEDLINE=21357209; PubMed=11463916;  
RA Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Rade D.,  
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
RA McDonald L.A., Feldblyum T.V., Argüoli S., Dickinson T., Hickey B.K.,  
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
RT "Complete genome sequence of a virulent isolate of Streptococcus  
RT pneumoniae";  
RL Science 293:498-506(2001).  
DR EMBL; AE007347; AAK74528.1; -.  
DR PIR; G95041; G95041.  
DR TIGR; SP0355; -.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005840; C:ribosome; IEA.  
DR GO; GO:0006189; F:methyltransferase activity; IEA.  
DR GO; GO:0003735; P:structural constituent of ribosome; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR GO; GO:0006412; P:protein biosynthesis; IEA.  
DR InterPro; IPR000878; Cor/por Metransf.  
DR InterPro; IPR007345; PS\_pyruv trans.  
DR Pfam; PF04230; PS\_pyruv trans; 1.  
DR PROSITE; PS00962; RIBOSOMAL\_S2\_1.  
KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 360 AA; 41405 MW; EC03E7E13AC357F2 CRC64;  
  
Query Match 7.8%; Score 105; DB 16; Length 360;  
Best Local Similarity 24.7%; Pred. No. 1.6;  
Matches 55; Conservative 34; Mismatches 76; Indels 58; Gaps 11;  
  
QY 31 IIVGLIERAKF-----PSN-SHTYDEKIAIQGFCGSLVEFVVVNTAIOAKAGRYKXP 83  
DB 26 IVDLLESKGYEVSFPNPSDFQMYLKEYRQKQ-----SFTKQADAILYIPGGYFGE 78  
QY 84 EENAFFPENLPPSIVPSYFQFLHPGAASININKSIWKYFKELPLLATSGDDGNVY 143  
DB 79 GHNARFRDNL-----IQKRFPLG-----IWASYFKKPIGVLGAGPN----- 118  
QY 144 TAANDLSLQISRIHYGKFAEVRKFD-----AP--QDYEPILIRAKDKEGLM 190  
DB 119 ---NDSLWNYGKRIINHAQFIT--VRDRESFDSLKHLSPSPVHETFDLIISKLREKT 174  
QY 191 KLTTFTSVEETVRKRV-----EKKAV-VFGQEVNLSDDNDN 226  
DB 175 EQLCOLKREAKDKKIILVYHNHKKALEKFAESISLFLNPN 217  
  
RESULT 15  
Q8PUQS ID Q8PUQS PRELIMINARY; PRT; 1259 AA.  
AC Q8PUQS;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein MW2277 (EC 2.7.3.-).  
GN MW2277.  
OS Methanosarcina mazei (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=22120827; PubMed=12125824;  
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
RA Martines-Arias R., Henne A., Wietzer A., Baeumer S., Jacobi C.,  
RA Brueggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,  
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RT "The genome of Methanosarcina mazei: evidence for lateral gene  
RT transfer between Bacteria and Archaea";  
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
DR EMBL; AE013469; AAM31973.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0016301; F:kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.  
DR GO; GO:0000160; P:two-component signal transduction system (p. .; IEA.  
DR InterPro; IPR003594; ATPbind ATPase.  
DR InterPro; IPR005467; His kinase.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000700; PAS-associat.  
DR InterPro; IPR000014; PAS\_domain.  
DR Pfam; PF02518; HATPase\_C; 1.  
DR Pfam; PF00785; PAC; 5.  
DR Pfam; PF00989; PAS; 1.  
DR TIGRFAMs; TIGR00229; sensory box; 6.  
DR PROSITE; PS50109; HIS\_KIN; 1.  
DR PROSITE; PS50113; PAC; 5.  
DR PROSITE; PS50112; PAC; 3.  
DR Kinase; Transferase; Complete proteome; Hypothetical protein.  
SQ SEQUENCE 1259 AA; 145131 MW; 9F523C6BF7A78440 CRC64;  
  
Query Match 7.4%; Score 98.5; DB 17; Length 1259;  
Best Local Similarity 25.9%; Pred. No. 29;  
Matches 59; Conservative 36; Mismatches 86; Indels 47; Gaps 15;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 22:50:31 ; Search time 4165 seconds  
(without alignments)  
10614.621 Million cell updates/sec

Title: US-10-624-061-15  
Perfect score: 1020  
Sequence: 1 gcacgagaggttaatttcaa.....atcttcaaaaaaaaaa 1020

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hcg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pt.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1020	100.0	1020	6	AR404688	AR404688 Sequence
2	465.6	45.6	525	6	AR404683	AR404683 Sequence
3	327.2	32.1	829	8	AY133840	AY133840 Arabidops
4	327.2	32.1	1006	6	AR236634	AR236634 Sequence
5	327.2	32.1	1006	6	AR236640	AR236640 Sequence
6	327.2	32.1	1006	6	AX343933	AX343933 Sequence
7	327.2	32.1	1006	6	AX464575	AX464575 Sequence
8	327.2	32.1	1006	6	AX464581	AX464581 Sequence
9	327.2	32.1	1006	8	L47355	L47355 Arabidopsis
10	327.2	32.1	1055	8	AY065238	AY065238 Arabidops
11	326.2	32.0	798	6	AX506277	AX506277 Sequence
12	326.2	32.0	993	6	AX343935	AX343935 Sequence
13	297.8	29.2	1000	8	L47356	L47356 Lycopersico
14	257.4	25.2	1231	6	AR404687	AR404687 Sequence
15	257.4	25.2	1250	8	AK068983	AK068983 Oryza sat
16	254.8	25.0	1274	8	AY089156	AY089156 Arabidops
17	254.8	25.0	1207	6	AR236633	AR236633 Sequence
18	254.8	25.0	1207	6	AR236639	AR236639 Sequence
19	254.8	25.0	1207	6	AX464574	AX464574 Sequence
20	254.8	25.0	1207	6	AX464580	AX464580 Sequence
21	254.8	25.0	1207	8	ATCHMUT	Z26519 A.thaliana
22	251	24.6	1005	6	AX506042	AX506042 Sequence
23	239.4	23.5	1223	6	AR404686	AR404686 Sequence
24	232.4	22.8	1435	8	AK101220	AK101220 Oryza sat
25	229.4	22.5	780	6	AR404689	AR404689 Sequence
26	229.4	22.5	780	8	BT009499	BT009499 Triticum
27	227.8	22.3	1043	8	AK069725	AK069725 Oryza sat
28	226	22.2	1348	8	AK105512	AK105512 Oryza sat
29	213	20.9	1143	8	AK117860	AK117860 Arabidops
30	212.6	20.8	951	8	BT005306	BT005306 Arabidops
31	211.4	20.7	1217	6	AR236635	AR236635 Sequence
32	211.4	20.7	1217	6	AR236641	AR236641 Sequence
33	211.4	20.7	1217	6	AX464576	AX464576 Sequence
34	211.4	20.7	1217	6	AX464582	AX464582 Sequence
35	211.4	20.7	1217	6	AF131219	AF131219 Arabidops
36	206.2	20.2	579	6	AR404684	AR404684 Sequence
37	173.4	17.0	8162	2	AC128660	AC128660 Medicago
38	132.4	13.0	246	8	AF012867	AF012867 Petroselin
39	129.8	12.7	394	8	PAJ004916	PAJ004916 Prunus av
40	123.2	12.1	2059	8	YSCAR07A	M24517 Saccharomyc
41	123.2	12.1	43776	8	SC9499X	Z49219 S.cerevisia
42	123.2	12.1	165536	8	SCCHRXVI	Z71255 S.cerevisia
43	99.8	9.8	843	6	BD266599	BD266599 Nucleic a
44	99.8	9.8	843	6	AX043758	AX043758 Sequence
45	99.8	9.8	1652	8	AF204738	AF204738 Fichia an

ALIGNMENTS

RESULT 1	AR404688	1020 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	Sequence 15 from patent US 6627798.				
DEFINITION	Sequence 15 from patent US 6627798.				
ACCESSION	AR404688				
VERSION	AR404688.1	GI:40153356			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 1020)				
AUTHORS	Falco,S.C., Famodu,O.O. and Lee,J.-M.				
TITLE	Aromatic amino acid biosynthetic enzymes				
JOURNAL	Patent: US 6627798-A 15 30-SEP-2003;				
FEATURES	Location/Qualifiers				

```
source 1. .1020
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 100.0%; Score 1020; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 1.7e-228; Indels 0; Gaps 0;
Matches 1020; Conservative 0; Mismatches 0;

QY 1 GCACGAGAGCTTAATTTCAAGATGCGCCAAAGCAGACAGCAACAAAGTCTGATCTCGGAA 60
DB 1 GCACGAGAGCTTAATTTCAAGATGCGCCAAAGCAGACAGCAACAAAGTCTGATCTCGGAA 60
QY 61 TGTGTACAGCTAGCTCTGTGAGAGGATTTGGTTAGGCAAGGATACCATCATTTA 120
DB 61 TGTGTACAGCTAGCTCTGTGAGAGGATTTGGTTAGGCAAGGATACCATCATTTA 120
QY 121 TGGTCTCATTTAGAGAGCCCAAGTTCCCTAGCAATTTCTCACCTATGATGAAAAGTATGC 180
DB 121 TGGTCTCATTTAGAGAGCCCAAGTTCCCTAGCAATTTCTCACCTATGATGAAAAGTATGC 180
QY 181 TCAATCCAGGGTTTTTGGGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGCCAT 240
DB 181 TCAATCCAGGGTTTTTGGGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGCCAT 240
QY 241 TCAAGCTTAAGGCTGGAAGATACAAAACCCCTGAAGAAAGCGCTTCTTCCAGAAAATTT 300
DB 241 TCAAGCTTAAGGCTGGAAGATACAAAACCCCTGAAGAAAGCGCTTCTTCCAGAAAATTT 300
QY 301 ACCACCATCAATTTGTCATCTTACTCTTCAACAGTTTGTGATCTCTGCTGCTTC 360
DB 301 ACCACCATCAATTTGTCATCTTACTCTTCAACAGTTTGTGATCTCTGCTGCTTC 360
QY 361 AATTACATAAAGTCCATCTGGAATATGTTTCAAGAGTACTTCCATTTGCTTGC 420
DB 361 AATTACATAAAGTCCATCTGGAATATGTTTCAAGAGTACTTCCATTTGCTTGC 420
QY 421 TACTTCGGGTGATGATGGCAATATATGCGCAAACTGCAGCTTAATGACCTTTCAATTTGCA 480
DB 421 TACTTCGGGTGATGATGGCAATATATGCGCAAACTGCAGCTTAATGACCTTTCAATTTGCA 480
QY 481 GTCCATCTAGAGAGTCACTATGGAAGTTTGTAGCTGAGTGAATGAGGATTCAGGGATGC 540
DB 481 GTCCATCTAGAGAGTCACTATGGAAGTTTGTAGCTGAGTGAATGAGGATTCAGGGATGC 540
QY 541 TCCTCAAGACTACGAGCCTTTAATTCGAGCTAAGGATAAAGAGGATTTGATGAAATTTGT 600
DB 541 TCCTCAAGACTACGAGCCTTTAATTCGAGCTAAGGATAAAGAGGATTTGATGAAATTTGT 600
QY 601 GACATTTAAGGCTTGAAGACGCTGAGGAGAGAGTTGAAGAAGGCTGTGCTGTT 660
DB 601 GACATTTAAGGCTTGAAGACGCTGAGGAGAGAGTTGAAGAAGGCTGTGCTGTT 660
QY 661 TGGCAGGAAGTGAATCTTAAAGTGAATGACATGACATGACATGACATGACATGACATGACAT 720
DB 661 TGGCAGGAAGTGAATCTTAAAGTGAATGACATGACATGACATGACATGACATGACATGACAT 720
QY 721 ATCAGTGGCTTTAGCTGTGTAACAAAATTTGGGTGATACCTCTCACCAAGAGGTTCAAGT 780
DB 721 ATCAGTGGCTTTAGCTGTGTAACAAAATTTGGGTGATACCTCTCACCAAGAGGTTCAAGT 780
QY 781 TGAGTACCTCTCGCGCTGACAGGCAATTAAGTGAATGAGTGAATTTAGAGAA 840
DB 781 TGAGTACCTCTCGCGCTGACAGGCAATTAAGTGAATGAGTGAATTTAGAGAA 840
QY 841 TGAAGATGAATATGATTTGTTTCAATGATTAAGCTCTTAAAGTATCCTTTATTCG 900
DB 841 TGAAGATGAATATGATTTGTTTCAATGATTAAGCTCTTAAAGTATCCTTTATTCG 900
QY 901 CAACTTCATCTAGCTGTGATTCAGAAATTAATTTGTAGCTATAGATTTTACCTTTAT 960
DB 901 CAACTTCATCTAGCTGTGATTCAGAAATTAATTTGTAGCTATAGATTTTACCTTTAT 960

RESULT 2
AE404683
LOCUS AR404683 525 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 5 from patent US 6627798.
ACCESSION AR404683
VERSION AR404683.1 GI:40153351
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 525)
AUTHORS Falco,S.C., Farnodu,O.O. and Lee,J.-M.
TITLE Aromatic amino acid biosynthetic enzymes
JOURNAL Patent: US 6627798-A 5 30-SEP-2003;
FEATURES Location/Qualifiers
source 1..525
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 45.6%; Score 465.6; DB 6; Length 525;
Best Local Similarity 95.4%; Pred. No. 1.5e-98;
Matches 499; Conservative 0; Mismatches 22; Indels 2; Gaps 2;

QY 8 AGCTTAATTTCAAGATGCGCCAAAGCAGACAGCAACAAAGTCTGATCTCGGATGTTAC 67
DB 1 AGCTTAATTTCAAGATGCGCCAAAGCAGACAGCAACAAAGTCTGATCTCGGATGTTAC 60
QY 68 ACGTAGCTCTCTGTGAGAGAGGATTTGGTTAGGCAAGGATACCATCATTTATGCTCTC 127
DB 61 ACGTAGCTCTCTGTGAGAGAGGATTTGGTTAGGCAAGGATACCATCATTTATGCTCTC 120
QY 128 ATTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACCTATGATGAAAAGTATGCTCAATC 187
DB 121 ATTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACCTATGATGAAAAGTATGCTCAATC 180
QY 188 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTTCAAGCT 247
DB 181 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTTCAAGCT 240
QY 248 AAGCTCGAAGATACAAAACCCCTGAAGAAAGCGCTTCTTCCAGAAAATTTACCAACCA 307
DB 241 AAGCTCGAAGATACAAAACCCCTGAAGAAAGCGCTTCTTCCAGAAAATTTACCAACCA 300
QY 308 TCAATTTGCGCATCTTACTCTTCAACACAGTTTGTGATCTCTGCTGCTTCAATTAAC 367
DB 301 TCAATTTGCGCATCTTACTCTTCAACACAGTTTGTGATCTCTGCTGCTTCAATTAAC 360
QY 368 ATAAACAAAGTCCATCTCGAAAATTTGATTTCAAGAGAGTACTTCCATTTGCTTACTTCG 427
DB 361 ATAAACAAAGTCCATCTCGAAAATTTGATTTCAAGAGAGTACTTCCATTTGCTTACTTCG 420
QY 428 GGTGATGATGCACTATGCGCAAACTGCAGCTTAATGACCTTTTCAATTTATGCAAGTCCATC 487
DB 421 GGTGATGATGCACTATGCGCAAACTGCAGCTTAATGACCTTTTCAATTTATGCAAGTCCATC 479
QY 488 TCTAGAAGGATTTCACTATGGAAGATTTGTAGCTGAGGTGAAT 530
DB 480 TCTAGAAGGA-TCACTATGGAAGATTTGTAGCTGANGNGAAAT 521

RESULT 3
AY133840
LOCUS AY133840 829 bp mRNA linear PLN 18-SEP-2002
DEFINITION Arabidopsis thaliana clone U18739 putative chorismate mutase CM2
(AT5G10870) mRNA, complete cds.
ACCESSION AY133840
VERSION AY133840.1 GI:22136859
```



KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
COMMENT  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
COMMENT  
FEATURES  
source  
gene  
CDS  
3'UTR

FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 829)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,  
Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R.,  
Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G.,  
Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M.,  
Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P.,  
Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Open Reading Frame (ORF) Clones  
Unpublished  
2 (bases 1 to 829)  
Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W.,  
Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K.,  
Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H.,  
Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A.,  
Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M.,  
Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M.,  
Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and  
Theologis, A.  
Direct Submission  
Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.  
The Salk, Stanford, PGECC (SSP) Consortium members constructed and  
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,  
Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M.,  
Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C.,  
Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Chen, H.,  
Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Miranda, M.,  
Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and  
Theologis, A.  
Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally  
to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP  
/PGECC) contributed equally to this work as PIs.  
Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.  
Location/Qualifiers  
1..829  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="U18739"  
/ecotype="Columbia"  
/note="This clone is in PUNI 51."  
1..829  
/gene="At5g10870"  
1..798  
/gene="At5g10870"  
/codon\_start=1  
/evidence=experimental  
/product="putative chorismate mutase CM2"  
/protein\_id="tAM91774.1"  
/db\_xref="GI:22136860"  
/translation="MARVPESDSGSGCNVLSLDLRESLIRQEDTIVFLIERAKPP  
LNSPAPESRLDGSFSSLEFFVRETEIIQAKVGEVPEENPFLENIPHSVPT  
HKYPSALHPKLSVNINQIWDIYFKELLPLFVPGDDGYNPSTAADLACLQSLSR  
IHYGKFAVAFRPAODYPAIRAQDRLEMLKLTIFKEVMEVKKVQKKAETFGQE  
VKFNSGYGDESKKYYKVDPLASRIYGEWLPILTKLVEVEYLLRLRD"  
799..829

ORIGIN  
Query Match 32.1%; Score 327.2; DB 8; Length 829;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
/gene="At5g10870"  
QY 22 AATGCCCAAGCAGCAGACAAAGCTCTGTTCTGGGAATGTGTACACGCTAGCTTCTGT 81  
Db 6 AAGAGCTTCGGAATCGGATTCGGTTCTGGTTTCTCCAAATGATCTGAGTCTTGTAAAT 65  
QY 82 GAGAGAGATTTGGTTAGCAGAGAGATACCATCATTTATGGTCTCATGTAGAGAGCAA 141  
Db 66 CAGAGAACTGTTGATTAGCAGAGAGACACCATCTCTTCAGCTTGTATCGAGAGAGCTAA 125  
QY 142 GTTCCCTAGCAATCTCACCTATGATGAAGATGTCTCAAAATCCAGGTTTTTGTGG 201  
Db 126 GTTCCCACTCAATCTCTGCTTTCGAGGAATCTCGTTGTCTAGATCTCGAAGTTTCTC 185  
QY 202 CTCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGGCTTGGAGATA 261  
Db 186 TTCTCTCACTGAGTTTTCGTCAGAGAGACAGAAATCATCAAGCTAAGTAGGAGATA 245  
QY 262 CAAAAACCTGAGAAAAAGCCCTTCTCCAGAAAAATTTACCAATCAATTTGTGCCATC 321  
Db 246 TGAATACCCGGAAGAGAAATCTTCTCTCTGAGAACATCTCTCACTCGGTTTTTCTCTAC 305  
QY 322 TTACTCTCTCAACAGTTTTTTCATCTCTGCTGCTCTCAATTAACATAAACAAGTCCAT 381  
Db 306 GCACAAATATCCTCGGTTTTCACCCCTAAGCTCTATCTGTATACATTACAAACAAT 365  
QY 382 CTGGAATAATGTTTCAAGAGATTACTTCCATGTTGTTCTTCTGTTGTTGTTGTTGTTG 441  
Db 366 CTGGGATATTTACTTTAAAGAAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 425  
QY 442 CTATCGCAACCTGAGCTAATGAGCTTTTCAATTTGAGTCTCATCTCTAGAGGATTCA 501  
Db 426 CTATCATCACTGCTGCTGATGATCTGCTGTTTCAAGCTCTTTCAGAGAGATTCA 485  
QY 502 CTATGAAAAGTTTGTAGCTGAGGTGAAATTCAGGATGCTCTCAAGACTACGAGCTTT 561  
Db 486 CTACGGTAAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCACAAGATTACGAGCTGC 545  
QY 562 AATTCAGCTAGGATAAAGAGGATTTGATGAATTTGATGACATTTACAGGCTTTGAAGA 621  
Db 546 GATTCGCGCTAGGATAGAGGCTTTGATGAAGCTGTTGAGCTTTGAGAAAGTAAAGA 605  
QY 622 GACGGTGAAGAGAGATTGAAAGAGGCTGTGTTTGGCAGGAGTGAATCTTTAA 681  
Db 606 AATGGTTAAGAGAGAGTCCAGAGAAAGCAGTTCGACAAAGTAAATTTCAA 665  
QY 682 -----CACTGATGACATGACAAATGAAACCGTAAATTTGATCTCATCGCTCTAG 735  
Db 666 CTCGCTGATGCGATGAGAGTAAGAAGATATAAAGTGGATCCATTGCTGCTCTCG 725  
QY 736 CTGTGTACAAAATTTGGTGTATACCTCTCACCAGAGAGTTTCAAGTTGAGTCACTCTTGG 795  
Db 726 CATCTACGGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGAGTATCTTCTACG 785  
QY 796 CGCTCTAGACTGAA 809  
Db 786 TCGTCTCGATTGAA 799  
RESULT 4  
AR236634  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Sequence 3 from patent US 6465217.  
AR236634  
AR236634.1  
GI:27280720  
Unknown.  
Unknown.

Unclassified.  
RESULT 5  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS AR236640/c  
LOCUS AR236640  
DEFINITION Sequence 9 from patent US 6465217.  
TITLE AR236640  
ACCESSION AR236640  
VERSION AR236640.1 GI:27280726  
JOURNAL Unknown.  
FEATURES Location/Qualifiers  
source 1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
QY 22 AATGGCCAAAGCAGCAGCAAAAGTCTGATTCCTGGAAATGTGTACACGCTAGCTTCTGT 81  
DB 69 AAGAGTCTTCGAATCGGATTCGGGTTCTGGTTGTTCCAAATGACTGAGTCTTGACTTAAT 128  
QY 82 GAGAGAGGATTGGTTAGGCAAGAGATACCATTTATGGTCTCATTTGAGAGAGCCAA 141  
DB 129 CAGAGAAATCGTTGATTAGGCAAGAGACCATTCGCTTCAGCTTGATCGAGAGAGCTAA 188  
QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAATCCAGGGTTTTTGTGG 201  
DB 189 GTTCCCACTCAATTTCTGCTTTCGAGGAATCTCGTTGCTAGATTCTGGAAGTTTCTC 248  
QY 202 CTCAATGGTGAATTTGTTTAAAGATACAGAGGCGCAATTCAGCTTAAGCTTGGAGATA 261  
DB 249 TTCTCTCACTGAGTTTTCGTAGAGAGACAGAAATCATCCAGCTTAAGGTAGGAGATA 308  
QY 262 CAAAAACCCCTGAAGAAACCCCTTCTTCCAGAAAAATTTACCAACCATCAATTTGTGCCATC 321  
DB 309 TGAATACCCGGAAGAGATCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 368  
QY 322 TTACTCTTCAACAGTTTTCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381  
DB 369 GCACAAATATPCCATCGCTTTCGACCTTAAGGCTTCTATCTGTTAAATTAACAAACAAT 428  
QY 382 CTGGAATAATGATTTCAAGAGTTACTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441  
DB 429 CTGGGATATTTACTTTAAGAAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 488  
QY 442 CTATGCGCAAACTGCAGCTAAATGACCTTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCT 501  
DB 489 CTATCCATCAACTGCTGCTAGTATCTCGCTGCTTTCACAAAGCTTCTTTCGAGAGGATTC 548  
QY 502 CTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCCTCAAGACTACGAGCCTTT 561  
DB 549 CTACGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTCCAGAGATTACGAGCCTGC 608  
QY 562 AATTCGAGCTAAGGATAAAGAGATTGATGAAATTTGTCATTTACAAAGCTTTGAAGA 621  
DB 609 GATTCGCGCTCAGGATAGAGGCTTTGATGAAGCTTTTGAAGCTTTTGAAGAGTAGAAGA 668  
QY 622 GACGGTGAAGAGAGATTGAAAGAGGCTGTGGTTTGGGCGAGGAGTGAATCTTAA 681  
DB 669 AATGGTTAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTTTGGACAGAGATTAATTCAA 728  
QY 682 -----CAGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
DB 729 CTCGCTATGGGATGAGAGTAAGAAAGATTAAGTGGATCCATTGCTTGCCTCTCG 788  
QY 736 CTGTGTAACAAATTTGGTATACCTTCTCACCAGAGGTTTCAAGTTGAGTACCTCTTGG 795  
DB 789 CATCTACGGGAGATGCTTATCCCTCTCACTAAGCTCGTTGAGGTGATATCTTCTACG 848  
QY 796 CCGCTCTAGACTGAA 809  
DB 849 TCGTCTCGATTGAA 862

Unclassified.  
RESULT 5  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS AR236640/c  
LOCUS AR236640  
DEFINITION Sequence 9 from patent US 6465217.  
TITLE AR236640  
ACCESSION AR236640  
VERSION AR236640.1 GI:27280726  
JOURNAL Unknown.  
FEATURES Location/Qualifiers  
source 1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
QY 22 AATGGCCAAAGCAGCAGCAAAAGTCTGATTCCTGGAAATGTGTACACGCTAGCTTCTGT 81  
DB 938 AAGAGTCTTCGAATCGGATTCGGGTTCTGGTTGTTCCAAATGACTGAGTCTTGACTTAAT 879  
QY 82 GAGAGAGGATTGGTTAGGCAAGAGATACCATTTATGGTCTCATTTGAGAGAGCCAA 141  
DB 878 CAGAGAAATCGTTGATTAGGCAAGAGACCATTCGCTTTCAGCTTGATCGAGAGAGCTAA 819  
QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAATCCAGGGTTTTTGTGG 201  
DB 818 GTTTCCTCACTCAATTTCTCTGCTTTCGAGGAATCTCGTTGCTAGATTCTGGAAGTTTCTC 759  
QY 202 CTCAATGGTGGAAATTTGTTTAAAGAAATACAGAGGCGCAATTCAGAGCTTAAGCTTGAAGATA 261  
DB 758 TTCTCTCACTGAGTTTTCGTAGAGAGACAGAAATCATCCAGCTTAAGGTAGGAGATA 699  
QY 262 CAAAAACCCCTGAAGAAACCCCTTCTTCCAGAAAAATTTACCAACCATCAATTTGCGCATC 321  
DB 698 TGAATACCCGGAAGAGAAATCTTCTTCTTGTGAGAAATCTCTCACTCGGTTTTTCTTAC 639  
QY 322 TTACTCTTCAAAACAGTTTTCGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 381  
DB 638 GCACAAATATCCATCGGCTTTGCAACCTAAGGCTCTATCTGTTAAATTAACAAACAAT 579  
QY 382 CTGGAATAATGATTTCAAGAGTTACTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 441  
DB 578 CTGGATATTTACTTTAAAGAAATTTGCTTCTTGTGTTGCAACCTCGGCGATGATCGCAA 519  
QY 442 CTATGCGCAAACTGCAGCTAAATGACCTTTCAATTTATGTCAGTCCATCTCTAGAGAGATTCA 501  
DB 518 CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAGCTCTTTCGAGAGGATTCA 459  
QY 502 CTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCCTCAAGACTACGAGCCTTT 561  
DB 458 CTACGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTCCCAAGATTTACGAGCTGC 399  
QY 562 AATTCGAGCTAAGGATAAAGAGATTGATGAAATTTGTCATTTTACAAAGCTTGAAGA 621  
DB 398 GATTCGCGCTCAGGATAGAGAGCTTTGATGAGCTTTTGAAGCTTTGAGAAAGTAGAAGA 339  
QY 622 GACGGTGAAGAGAGATTGAAAGAGGCTGTGGTTTGGGCGAGGAGTGAATCTTAA 681  
DB 338 AATGGTTAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTTTGGACAGAGATTAATTCAA 279  
QY 682 -----CAGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735

```
Db 278 CTCTGGCTATGGCGATGAGATGAAGAAAGATATAAGTGGATCCATTGCTTGCCTCTCG 219
Qy 736 CTGTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTACCTCTTGGC 795
Db 218 CATCTACGGGAATGGCTTATCCCTCTCACTAAAGCTCGTTGAGGTTGAGTATCTTCTACG 159
Qy 796 CCGTCTAGACTGAA 809:
Db 158 TCGTCTCGATTGAA 145

RESULT 6
AX343933
LOCUS AX343933 1006 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 3 from Patent WO200901.
ACCESSION AX343933
VERSION AX343933.1 GI:18491972
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Badur, R., Geiger, M., Kunze, I. and Sommer, S.
TITLE Changing the fine chemical content in organisms by genetically
modifying the shikimate pathway
JOURNAL Patent: WO 020901-A 3 03-JAN-2002;
Sungene GmbH & Co. KGAA (DE)
FEATURES
source
1. .1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
64. .861
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22398.1"
/db_xref="GI:18491973"
/db_xref="REMBEL:CAD22398"
/translation="MARVFESDSSGNSVLSLDLRESLIREDTIVFSLIBRAKFP
KNSPAGESCLDSGSFSSLTFFVETELIQAKGRYEPYEPENPFLENPHSVFT
HYPFSLHKLSDINIKQWIDYIFKELLPFVKPGDGNYPSTADLACLQALRR
IHYGKFAVEKFRADPODYPAPRAODREALMKLLTPKEVEMVKRVKKAFTFQGE
VKFNSGYGDESCKKYKVDPLLASRIYGEWLIPLTKLVEVEYLLRLD"

CDS
ORIGIN
Query Match 32.1%; Score 327.2; DB 6; Length 1006;
Best Local Similarity 64.2%; Pred. No. 4e-66;
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;
Qy 22 AATGCCAAAGCAGACAGACAAAGTCTCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81
Db 69 AAGAGTCTTCAATCGGATTCGGGTTCTGGTTGTTCATGTACTGAGTCTTGACTTAAT 128
Qy 82 GAGAGAGATTGGTTAGGCAAGAGGATACCATATTATGTTCTCATTTGAGAGAGCCAA 141
Db 129 CAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGAGTAA 188
Qy 142 GTTCCCTAGCAATCTCACACTATGATGAAGATGCTCAATCCAGGTTTGTGG 201
Db 189 GTTCCACTCAATCTCTCTGTTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGTTTCTC 248
Qy 202 CTCATTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAAGCTAAGCGTGAAGATA 261
Db 249 TTCTCTCACTGAGTTTTCGTCAGAGAGACAGAAATCATCCAACTAAGCTAGGAAGATA 308
Qy 262 CAAAMACCTGAAGAAAGCGCTTCTCCAGAAATTTACCACCATCAATTGTGCCATC 321
Db 309 TGAATACCCGGAAGAGATCTCTTCTTCCTTGAGAACATTCCTCAGCTGGTTTTCCTAC 368
Qy 322 TTACTCTTCAACAGATTTCGATCTCTGCTGCTTCAATTAACATAACAGAGTCCAT 381
```

```
Db 369 GCACAAATATCCATCGGCTTTGCACCTAAGGCTCTATCTGTAAATTAACAAACAAT 428
Qy 382 CTGGAAAATGTATTTCAAAGAGTTACTTCCATTGCTTGTCTACTTCGGGTGATGATGGCAA 441
Db 429 CTGGGATATTTACTTTAAAGAAATGCTTCCCTTTGTTTGTCAAACCTGGCGATGATGGCAA 488
Qy 442 CTATGCGCAAACCTGCGAGCTAATGAGCTTTTCATTATTGCAAGTCCATCTCTAGAGGATTCA 501
Db 489 CTATCCATCAACTGCTGCTAGTATGATCTCGCTGTTTACAAGCTCTTTCGAGAGGATTCA 548
Qy 502 CTATGGAAGATTTGTAGCTGAGTGAGTGAATTCAGGAGATGCTCTCAAGACTACGAGCTTT 561
Db 549 CTACGGTAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAGATTACGAGCTGTC 608
Qy 562 AATTGAGCTAAGGATAAAGAGATTGATGAAATTTGATGACATTTACAAGCGTTGAAGA 621
Db 609 GATTCCGCTCAGGATAGAGAGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGAAGA 668
Qy 622 GACGCTGAGGAAGAGATTGAAGAGAGCTGTGCTGTTTGGGCGAGGAGTCAATCTTAA 681
Db 669 AATGGTTAAGAGAGAGTGCAGAGAAAGACAAAGCTTTGGACAAAGATGAATTAATCAA 728
Qy 682 -----CACTGATGACAATGACAATCAAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735
Db 729 CTCTGCTATGCGATGAGAGTAAGAAGATATAAGTGGATCCATTGCTTGCCTCTCG 788
Qy 736 CTGTACAAAATTTGGTGATACCTCTCACCAAGAGGTTTCAAGTTGAGTACTCTTGGC 795
Db 789 CATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG 848
Qy 796 CCGTCTAGACTGAA 809
Db 849 TCGTCTCGATTGAA 862

RESULT 7
AX464575
LOCUS AX464575 1006 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 3 from Patent WO202798.
ACCESSION AX464575
VERSION AX464575.1 GI:21899370
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Boyes, D.C., Davis, K.R., Woessner, J.P., Gorlach, J., Hamilton, C.M.,
Hoffman, N.B., Kloti, A.S., Zayed, A., Ascenzi, R.A., Allen, K.,
Mulpuri, R. and Kjemtrup, S.
TITLE Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: WO 0202798-A 3 10-JAN-2002;
Paradigm Genetics Inc. (US)
FEATURES
source
1. .1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 32.1%; Score 327.2; DB 6; Length 1006;
Best Local Similarity 64.2%; Pred. No. 4e-66;
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;
Qy 22 AATGCCAAAGCAGACAGACAAAGTCTCTGGAATGTGTACACGCTAGCTTCTGT 81
Db 69 AAGAGTCTTCAATCGGATTCGGGTTCTGGTTGTTCATGTACTGAGTCTTGACTTAAT 128
Qy 82 GAGAGAGATTGGTTAGGCAAGAGGATACCATATTATGTTCTCATTTGAGAGAGCCAA 141
Db 129 CAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTTCAAGCTCGTTGAGTATCTTCTACG 188
```

```

QY 142 GTTCCCTAGCAATTCACACCTATGATGAAAGTATGCTCAAAATCCAGGGTTTTGTGG 201
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 GTTCCCACTCAATTCCTCTCTCGAGGAATCTGTTGTCTAGATTCTGGAAGTTTCTC 248

QY 202 CTCAATGGTGAATTTGTTTGAAGATACAGAGGCATTCAGCTTAAGCTGGAAGATA 261
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 TTCTCTCACTGAGTTTTGCTCAGAGAGACAGAAATCATCCAAGCTTAAGGTAGGAATA 308

QY 262 CAAAAACCCCTGAAGAAACCCCTCTTCCAGAAAAATTTACCACCAATCAATTTGTGCCATC 321
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 TGAATACCGGAGAGATCTTTCTTCTTGAGAACATTCCTCACTCGTTTTTCTTAC 368

QY 322 TTACTCTTCAACAGTTTTTGATCTCTGGTGTCTGCTTCAATTAAACATAAACAGTCCAT 381
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 369 GCACAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTTAACATTAACAAACAAT 428

QY 382 CTGMAAATGATTTCAAGAGTTACTTCCATTGCTTTGTACTCTCGGTTGATGGCAA 441
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 429 CTGGATATTTACTTTAAAGATTTGCTTCTTTGTTTGTCAACCTGGCGATGGCAA 488

QY 442 CTATGCAAACTGACGCTTAATGACCTTTCAATTAATTCAGTCCATCTCAGAGGATCA 501
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 489 CTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAAGTCTTTTCGAGAGGATTC 548

QY 502 CTATGGAAGTTGTAGCTGAGGTGAATTCAGGGATGCTCTCAAGACTACGAGCCTTT 561
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 549 CTAGGTAAATTTGTAGCTGAGTCAATTCAGAGATGCTTCCACAGATTCAGAGCTGC 608

QY 562 AATTGCAAGCTAAGGATAAAGAGGATTTGATGAATTTGTGACATTTACAAGGTTGAAGA 621
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 609 GATTGCGCTCAGGATAGAGGCTTTGATGAAGCTTTTGACGTTTGAGAAAGTAGAAGA 668

QY 622 GAGGTGAGGAAGAGATTGAAGAGGCTGTGGTTTGGCAGGAGTGAATCTTAA 681
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 669 AATGGTTAAGAGAGAGTGAGAGAGAAAGCAGAAACGTTTGGACAGAAAGTAGAATA 728

QY 682 -----CAGTGATGACAAATGACAAATGAAACCCGTAATTTGATCCATCAGTGGCTCTAG 735
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 729 CTCTGGCTATGGCGATCAGAGTAGAAGAGATATAAGTGGATTCATTGCTTGCCTCTCG 788

QY 736 CTGTACAAAATTTGGTGATCTCTCACCAGAGGTTGAGTTGATGATCTCTTGGG 795
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 789 CATCTAGGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTAGG 848

QY 796 CCGTCTAGACTGAA 809
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 849 TCGTCTCGATTGAA 862

RESULT 8
AX464581/c
LOCUS AX464581
DEFINITION Sequence 9 from Patent WO0202798.
ACCESSION AX464581
VERSION AX464581.1 GI:21899376
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Boyes, D.C.; Davis, K.R.; Woessner, J.P.; Gorlach, J.; Hamilton, C.M.;
Hoffman, N.E.; Kloti, A.S.; Zayed, A.; Ascenzi, R.A.; Allen, K.;
Mulpuri, R. and Kjettrup, S.
METHODS Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: WO 0202798-A 9 10-JAN-2002;
Paradigm Genetics Inc. (US)
FEATURES
source Location/Qualifiers
1..1006
/organism="Arabidopsis thaliana"

```

```

/mol_type="unassigned DNA"
/db_xref="taxon:3702"

Query Match 32.1%; Score 327.2; DB 6; Length 1006;
Best Local Similarity 64.2%; Pred. No. 4e-66;
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

ORIGIN
QY 22 AATGCCAAAGCAGACAGACAAAGTCTCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 938 AAGAGTCTTCGAATCGGATTCTGGTCTGCTGTTGTTCCATGTACTGAGTCTTGACTTAAT 879

QY 82 GAGAGAGATTTGGTTAGGCAAGAGGATACCATCATTTATGTCTCATTTGAGAGAGCCAA 141
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 878 CAGAGAAATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGAATCAGAGAGCTAA 819

QY 142 GTTCCCTAGCAATTCACACCTATGATGAAAGTATGCTCAAAATCCAGGGTTTTGTGG 201
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 818 GTTTCACACTCAATTTCTCTGCTTTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGTTTCTC 759

QY 202 CTCAATGGTGAATTTGTTTGAAGATACAGAGGCATTCAGCTTAAGCTGGAAGATA 261
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 758 TTCTCTCACTGAGTTTTTGTCTCAGAGAGACAGAAATCATCCAAGCTAAGGTAGGAATA 699

QY 262 CAAAAACCCCTGAAGAAACCCCTCTTCCAGAAAAATTTACCACCAATCAATTTGTGCCATC 321
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 698 TGAATACCCGAGAGAGATCTTTCTTCTTGAGAACATTCCTCACTCGTTTTTCTTAC 639

QY 322 TTACTCTTCAACAGTTTTTGATCTCTGGTGTCTGCTTCAATTAACATAAACAGTCCAT 381
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 638 GCACAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTTTAACATTAACAAACAAT 579

QY 382 CTGGAATAATGATTTCAAGAGTTACTTCCATTGCTTGTCTACTCTCGGTGATGAGCAA 441
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 578 CTGGATATTTACTTTAAAGANTTGTCTTCTTTGTTGTTCACACCTGGCGATGAGCAA 519

QY 442 CTATGCCAAACTGCAAGCTAATGACCTTTCAATTTATGTCAGTCCATCTCTAGAAGGATCA 501
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 518 CTATCCATCACTGCTGCTAGTATCTCGCTCTTTTACAAGCTCTTTTCGAGAGAGGATCA 459

QY 502 CTATGGAAAGCTTTGATGCTGAGGTGAATTCAGGATGCTCTCAAGACTACGAGCCTTT 561
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 458 CTACGGTAAATTTGTAGCTGAGGTCAATTCAGAGATGCTTCCACAAGATTTACAGGCTGC 399

QY 562 AATTGCAAGCTAAGGATAAAGAGGATTTGATGAATTTGTGACATTTACAAGGCTTGAAGA 621
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 GATTGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACCGTTTGAGAAAGTAGAAGA 339

QY 622 GAGGTGAGGAAGAGATTGAAGAGGCTGTGGTTTGGCAGGAGTGAATCTTAA 681
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 AATGGTTAAGAGAGAGTGAGAGAGAAAGCAGAAACGTTTGGACAGAGATTAATTTCAA 279

QY 682 -----CAGTGATGACAAATGACAAATGAAACCCGTAATTTGATCCATCAGTGGCTCTAG 735
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 CTCTGGCTATGGCGATCAGAGTAGAAGAGATATAAGTGGATTCATTGCTTGCCTCTCG 219

QY 736 CTGTACAAAATTTGGTGATCTCTCACCAGAGGTTGAGTTGATGATCTCTTGGG 795
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 CATCTAGGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTAGG 159

QY 796 CCGTCTAGACTGAA 809
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 TCGTCTCGATTGAA 145

RESULT 9
LOCUS L47355
DEFINITION Arabidopsis thaliana chorismate mutase mRNA, complete cds.
ACCESSION L47355
VERSION L47355.1 GI:5732015
KEYWORDS chorismate mutase; mutase.
SOURCE Arabidopsis thaliana (thale cress)

```

ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS Eberhard, J., Ehler, T.T., Epple, P., Felix, G., Raesecke, H.R.,  
Amrhein, N. and Schmid, J.  
TITLE Cytosolic and plastidic chorismate mutase isozymes from Arabidopsis  
thaliana: molecular characterization and enzymatic properties  
JOURNAL Plant J. 10 (5), 815-821 (1996)  
MEDLINE 97111372  
PUBMED 8953244  
COMMENT GSDS:S:46595  
FEATURES  
Location/Qualifiers  
1..1006  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/cldones="catCM2"  
64..861  
/EC\_number="5.4.99.5"  
/note="putative"  
/codon\_start=1  
/product="chorismate mutase"  
/protein\_id="AAD48922.1"  
/db\_xref="GI:5732016"  
/translation="MARVPSDSGSCSNVSLDLRESLIRQEDTIVPSLIERAKFP  
LNPAFESCLSGSSSLTEFVRETEIIQAKVGRYEPENPFLENPHSVFPT  
HKYPSALHPKALSNINKIWDIYFKLEFLFKPDGDNYPSTASDLACIQAUSRR  
HFGKFAVEYKFDPAQDREALMKLLTFKVEEMVKRVKQKATFQGE  
VFNKSGYGSDESKKKYKVDPLLASRIYGEWLIPLTKLVEVEYLLRLD"  
1006  
polya\_site /note="putative"  
ORIGIN  
Query Match 32.1%; Score 327.2; DB 8; Length 1006;  
Best Local Similarity 64.2%; Pred. No. 4e-66;  
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
Qy 22 AATGCCAAAGCAGCAGCAAAAGTCTCTGATTCGGGAAGTGTACACGCTAGCTCTGT 81  
Db 69 AAGAGTCTTCAATCGGAATCGGGTCTCGTGTTCCTCAATGACTAGTCTTCACTTAAT 128  
Qy 82 GAGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGGTCTCATTCAGAGAGCCAA 141  
Db 129 CAGAGAATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGAGCTAA 188  
Qy 142 GTTCCCTAGCAATCTCACACCTATGATGAAGATGCTCAATCCAGGCTTTTGTGG 201  
Db 189 GTTTCCTCAATCTCTCTGTTTCGAGGAATCTCGTGTCTAGATTCGGAAGTTCTC 248  
Qy 202 CTCATTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAGCTAAGGCTGGAAGATA 261  
Db 249 TTCTCTCACTGAGTTTTCGTACAGAGACAGAATCATCCAGCTAAGGTAGAGATA 308  
Qy 262 CAAAAACCTGAAGAAACGCTTTCTTCCAGAAAAATTTACCAACCATCAATTGTGCAATC 321  
Db 309 TGAATACCGGAGAGAAATCTTTCTTCTTCTTGAAGAAATCTCTCACTCGGTTTTTCTAC 368  
Qy 322 TTACTCTTCAACAGATTGTCATCTGCTGCTTCAATTAACATAACCAAGTCCAT 381  
Db 369 GCACAAATATCCATCGGCTTTCACCTTAAGCTCTATCTGTTACATTAACAAACAAT 428  
Qy 382 CTGGAAGATGTTTCAAGAGTTACTTCCATTCCTTCTGCTTCTCGGGTGATGATGCCAA 441  
Db 429 CTGGATATTACTTTAAAGAAATGCTTCTCTTCTTGTCTCAACCTCGGATGATGCCAA 488  
Qy 442 CTATGCGCAACTGACGTAATGACCTTTCATTTATTCAGTCCATCTCTAGAGAGTTCA 501  
Db 489 CTATCCATCAACTCTGCTAGTATCTCGCTGTTTCAAGCTCTTTTCGAGAGAGGATCA 548  
Qy 502 CTATGGAAGTTTCTAGTCTGAGTGAAATTCAGGATGCTCTCAAGACTACAGGCTTT 561

Db 549 CTACGGTAAATTTAGCTGAGGTCAAATTCAGAGATCTCCACAAGATTACGACCTGC 608  
Qy 562 AATTCAGCTAGGATAAAGAGGATTTGATGAATTTGTTGACATTTACAAGCTTTGAAGA 621  
Db 609 GATTCGGCTCAGGATACAGAGGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGAAGA 668  
Qy 622 GACGTTGAGGAGAGAGTTGAAAGAGAGCTGTGTTGTTGGGAGGAGTGAATCTTAA 681  
Db 669 AATGTTAAGAGAGAGTGTCAGAGAAAGACGAAACGTTTGACAGAAAGTAAATCAA 728  
Qy 682 -----CAGTGATGACAATGACAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
Db 729 CTCTGCTATGCGGATGAGAGTAAGAAGATATAAAGTGATCCATTTGCTTGCCTCG 788  
Qy 736 CTTGTACAAAATTTGGGTGATACCTTCCACAGAGAGTTGAGTTGAGTACTCTTGG 795  
Db 789 CATCTACGGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGATCTTCTACG 848  
Qy 796 CCGTCTAGACTGAA 809  
Db 849 TCGTCTGATTTGAA 862  
AY065238 1055 bp mRNA linear PLN 18-SEP-2002  
Arabidopsis thaliana putative chorismate mutase CM2 (At5g10870)  
mRNA, complete cds.  
AY065238  
AY065238.1 GI:17529007  
FLI CDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 1055)  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full Length cDNA Clones  
Unpublished  
2 (bases 1 to 1055)  
Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M.,  
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,  
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,  
Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A.,  
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN  
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.  
The Saik, Stanford, PGEC (SSP) Consortium members constructed and  
sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K.,  
Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M.,  
Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C.,  
Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T.,  
Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M.,  
Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PFGE) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinohara, K. (RIKEN GSC) and Theologis, A. (SSP/PFGE) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

## FEATURES

## source

Location/Qualifiers

1. 1055

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/chromosome="5"

/clone="RAF109-41-E09 (R18739)"

/ecotype="Columbia"

/notes="This clone is in a modified pBluescript vector (FUC-1) as a BamHI/XhoI insert."

1. 1055

/gene="At5g10870"

1. 98

/gene="At5g10870"

2

misc\_difference

/gene="At5g10870"

/notes="compared to genomic sequence"

/replaces="t"

99. 896

/gene="At5g10870"

/codon\_start=1

/evidence=experimental

/product="putative chorismate mutase Cm2"

/protein\_id="AAL38714.1"

/db\_xref="GI:17529008"

/translation="MARVPSDSGSGSNVLSLIRSELIROEDTIVFSLIERAKPP

LSNPARESRCLDSGFSLSLTFVRETEIIQAKVGRVYEPENFFLENPHSVFPT

HKYPSALHPKALSVNINKQDIYFKLELLPLFKPDGDDNPSTPAASDLAQLSRR

IHYGKFAVEKPDQDEYPAIRAQDREALMLLTFEKVEEMVKRQKAEFTGQE

VFNISGYSDESKKKYKVDPLLASRIYGEWLIPLTKLIVEVEYLLRLD"

897. 1055

/gene="At5g10870"

## 3'UTR

## ORIGIN

Query Match 32.1%; Score 327.2; DB 8; Length 1055;

Best Local Similarity 64.2%; Pred. No. 4e-66;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCCAAAGCAGCAGCAACAAAGTCCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81

DB 104 AAGAGTCTCGAATCGAATCGGGTCTGGTGTCTCAATGACTGAGTCTTGACTTAAT 163

QY 82 GAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGGTCTCAATGAGAGGCCAA 141

DB 164 CAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGAGCTAA 223

QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAGATGCTCAATCCAGGTTTGTGG 201

DB 224 GTTCCACTCAATCTCCGCTTCAGGAATCTCGTGTCTAGATTCGGAAGTTCTC 283

QY 202 CTCATTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAAGCTTAAGCTCGAAGATA 261

DB 284 TTCTCTCAGTGATTTTGTGTCAGAGACAGAAATCATCAAGCTTAAGTAGGAAGATA 343

QY 262 CAAAACCTTGAGAAAACGCTTCTTCCAGAAAATTTACCAACCATCAATTTGTCGCATC 321

DB 344 TGAATACCCGGAAGAGAATCTTCTTCTCTTGAGAACATTCCTCACTCGGTTTTTCCCTAC 403

QY 322 TTACTCTCTCAACAGTTTTTTCATCTCGTGTCTCAATTAACATAAACAAGTCCAT 381

DB 404 GCACAAATATCCATCGGCTTTCACCTTAAGGCTCTATCTGTAAATTAACAACAAT 463

QY 382 CTGGAAAATGATTTCAAGAGTTACTTCGATGCTTGCTTACTTCCGGTGATGATGGCAA 441

DB 464 CTGGGATATTACTTTTAAAGAAATGCTTCTCTTTGTTGTCAAACTGGCGATGGCAA 523

QY 442 CTATGGCAAACTGCAGCTAATGACCTTTTCATTTATTTGCAGTCCATCTCTAGGAAGATTCA 501

DB 524 CTATCCATCAACTCTCTAGTATCTCGGCTGTGTACAAAGCTCTTTTCGAGAAGGATTCA 583

QY 502 CTATGGAAAGTTTCTAGCTGAGTGAATTCAGGATGCTCCTCAAGACTTACGAGCTTT 561

DB 584 CTACGGTAAATTTCTAGCTGAGTCAAAATTCAGATGCTCCCAAGATTACGAGCTGC 643

QY 562 AATTCGAGCTAAGGATAAAGAGGATTGATGAATTTGTTGACATTTACAAGCTTTGAAGA 621

DB 644 GATTCCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTCAGAAAAGTAGAAGA 703

QY 622 GACGTTGAGGAGAGAGTTGAAAAGAGGCTGCTGTTGTTGGCAGGAGTGAATCTTAA 681

DB 704 AATGGTTAAAGAGAGAGTGCAGAGAAGCAGAAACGTTTGGACAGAGATAAATTCAA 763

QY 682 -----CAGTGATGACCAATGCAATGAAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735

DB 764 CTCTGGCTATGGCGATGAGAGTGAAGAGATATAAAGTGATGATCCATTCCTCTCTCG 823

QY 736 CTCTACAAAATTTGGGTGATACCTCTCACCAGAGGTTTCAGTTGAGTACCTCTTGG 795

DB 824 CATCTACGGGAATGGCTTATCCCTCTCACTAAGCTGCTTGGAGTTGAGTATCTCTACG 893

QY 796 CCGTCTAGACTGAA 809

DB 884 TCGTCTCGATTGAA 897

## RESULT 11

AX506277

LOCUS

AX506277

DEFINITION

AX506277

ACCESSION

AX506277.1

VERSION

AX506277.1

KEYWORDS

ORGANISM

Arabidopsis thaliana (thale cress)

SOURCE

Arabidopsis thaliana

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 798

/organism="Arabidopsis thaliana"

/mol\_type="unassigned DNA"

/db\_xref="taxon:3702"

ORIGIN

Query Match 32.0%; Score 326.2; DB 6; Length 798;

Best Local Similarity 64.2%; Pred. No. 7e-66;

Matches 509; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCCAAAGCAGCAGCAACAAAGTCCTGATTCTGGGAATGTGTACACGCTAGCTTCTGT 81

DB 6 AAGAGTCTTTCGAATCGAATTCGGGTTCTGGTGTCTCAATGACTGAGTCTTGACTTAAT 65

QY 82 GAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGGCCAA 141

DB 66 CAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGTATCGAGAGACTAA 125

QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAGATGCTCAATCCAGGTTTTTGTGG 201

DB 126 GTTTCCTCAATCTCTCTGCTTCGAGGAATCTCGTGTCTAGATTCGGAAGTTCTC 185

QY 202 CTCATTGGTGAATTTGTTTAAAGATAACAGAGGCCATTCAGAGCTTGAAGATA 261

DB 186 TTCTCTCACTGAGTTTTTTCGTCAGAGAGACAGAAATCATCAAGCTTAAGTAGGAAGATA 245





```

ACCESSION L47356
VERSION L47356.1 GI:5732017
KEYWORDS choriolate mutase, mutase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Eberhard, J., Bischoff, M., Raesecke, H.R., Amrhein, N. and Schmid, J.
TITLE Isolation of a cDNA from tomato coding for an unregulated,
cytosolic choriolate mutase
JOURNAL Plant Mol. Biol. 31 (4), 917-922 (1996)
MEDLINE 96400046
PUBMED 8806422
COMMENT GSDB:S:46596
FEATURES
    Location/Qualifiers
        1..1000
            /organism="Lycopersicon esculentum"
            /mol_type="mRNA"
            /strain="UC82b"
            /db_xref="taxon:4081"
            /clone="cLeCM2"
            /EC_number="5.4.99.5"
            /note="putative"
            /codon_start=1
            /product="choriolate mutase"
            /protein_id="AAD48923.1"
            /db_xref="GI:5732018"
            /translation="MACGADKLCLDSIRKSLIRODTIIFNLIERIKFPINPTLYKQ
            OLPPSNRSGSLFOYLFORTESLOKVGRLAPEENPPFDNLSDSIIPLKCTPVLH
            BRAESVYNEKILDIYNOMLPLECTEVNDANEATTAACDIQLQLSRRIHYGKFV
            BEVKRDSIDEXKPFILAQDRDALMKLITFEAVEEMKRVAKAKVFGQEVSLDNA
            BEVKRSDLPVLSRLYDEWVMPVLTKEVEYVLRRLD"
        polyA_site
            1000
            /note="putative"
    ORIGIN
        Query Match 29.2%; Score 297.8; DB 8; Length 1000;
        Best Local Similarity 64.5%; Pred. No. 3.1e-59; Indels 6; Gaps 2;
        Matches 478; Conservative 0; Mismatches 257;
    QY 71 CTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGATACCATCATTTATGGTCTCAT 130
    DB 107 CTTGATTCAATAAGAAAGAGTTTAAATTAGACAAGAGATACCATTATTTCACCTTAT 166
    QY 131 GAGAGAGCCAGTTCCTAGCATTTCTCACCTATGATGAAGATGCTCAATCCAG 190
    DB 167 GAAAGAAATCAAAATTCCTCAATTAATCAATTTATACAAACAAATACCTCTCTTCC 226
    QY 191 GGTGTTTGTGGCTCATTTGGTGGAAATTTGTTTGAAGATACAGAGCCATTCAGCTAAG 250
    DB 227 AATTTTTCAGATCCTTATTTCAATCTGTTTTCAGAAACAGATCCCTTCAATCTAAG 286
    QY 251 GCTGGAAGATACAAAACCTGAGAAAAGCCCTTCTCCAGAAAATTTACCAACATCA 310
    DB 287 GTTGGTAGATATTTGGCCCGCAGAGAAAATCCCTTCTCCAGATAATTTGCTGAATCA 346
    QY 311 ATTGTGCTATCTTACTCTTCAACAGTTTGTGATCTGCTGCTGCTTCAATTAACATA 370
    DB 347 ATCATACCACTTACAAATGACACACAGTTTGGATCTTCTGAGCAGATCAGTAATGTA 406
    QY 371 AACAAAGTCCATCTGGAATAATGTAATTTCAAGAGTTACTTCCATTTGCTGCTACTTGG-- 428
    DB 407 AATGAGAAAATATTGGATATTTATATAAATCATGATGCTTCCACTATTTTGTACCGAGTT 466
    QY 429 -GTGATGATGGCAACTATGCGCAAACTGCGAGCTAATGACCTTTTCAATTTGCGAGTCCATC 487
    DB 467 AATGATGATGCAAAATTTTGGCACTACTGCTGCTTGTGACATTCAGCTATTGAGGCACTA 526
    QY 488 TCTAGAGAGATTCATATGAAAGTTTGTAGCTGAGGTGAATTCAGGATGCTCTCTCA 547

```

---

```

Db 527 TCTAGAAAGGATTCACTATGGAATAATTTGTTGCTGAAGTTAAATTTAGAGATAGATTGAT 586
QY 548 GACTTACGAGCCTTTAATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGTTGACATTT 607
Db 587 GAATATAAACCAATTTATTCTTCTCAGGATAGAGATGCTTTAATGAAGTTATTAAACATTT 646
QY 608 ACAAGCGTTGAAGAGACGCTGAGGAAGAGATTGAAAAGAGGCTGCTGCTGTTTGGCGAG 667
Db 647 GAAGCGGTTGAAGAGATGTTGGAAGAGAGATGAGCGAAGAGCTAAAGTTTGTCTAA 706
QY 668 GAAGTGAATCTTAACAGTGTATGACAAATGACAAATGAAAACCGTAAATTTGATCCATCAGTG 727
Db 707 GAAGTGAAGCTTAAC--GATAATGCTGAAGAGTGAAGGCAAAATTCATCCATTTGTA 763
QY 728 GCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCAGAGAGGTTGAGTGTAGTAC 787
Db 764 GTTTCACGTTTATATGATGAATGGTTATGCTTGTGCTTGTGCTGCTGCTGCTGCTGCT 823
QY 788 CTCTTGGCGCTCTAGACTGA 808
Db 824 CTCTGCTGCTCTTGTGATTA 844

RESULT 14
AR404687 LOCUS 1231 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 13 from patent US 6627798.
ACCESSION AR404687
VERSION AR404687.1 GI:40153355
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1231)
AUTHORS Falco, S.C., Farnodu, O.O. and Lee, J.-M.
TITLE Aromatic amino acid biosynthetic enzymes
JOURNAL Patent: US 6627798-A 13 30-SEP-2003;
FEATURES
    Location/Qualifiers
        1..1231
            /organism="unknown"
            /mol_type="genomic DNA"
    ORIGIN
        Query Match 25.2%; Score 257.4; DB 6; Length 1231;
        Best Local Similarity 58.4%; Pred. No. 8.9e-50;
        Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;
    QY 24 TGCCAAAGCAGCAGCAACAAAGTCTGATTCTGGGAATGTGTACACGCTAGCTTCTGTGA 83
    DB 239 TGACCAAGGAGGAGAAACAGAGGATAGATCAAAAGTGAATCTGACCTTGGACACATTA 298
    QY 84 GAGAGGATTTGGTTAGCAAGAGGATACCATCTATTATGGTCTCATTTGAGAGAGCCAGT 143
    DB 299 GAACCTCCTTGGTTAGCAAGAGAGAGATCATATTTACGCTCTTTAGAGAGAGACAGT 358
    QY 144 TCCTAGCATTTCTCACCTATGATGAAGATGCTCAAAATCCAGGGTGTGTTGGCT 203
    DB 359 TTGCTCAATGCTGATATATATGATAAATAATGCTTTCAATGGGATGATTTGATGGCT 418
    QY 204 CATTTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAGCTAAGCTGAGGAGATACA 263
    DB 419 CTTTGGTTGAATTCATGTTTAGAGAAACCGAAAACCTACATCAACAGGTTGGAGATACA 478
    QY 264 AAAACCTGAGAAAGCGCTTCTCCGAAAATTTACCAATCAATTTGTCCTATCTT 323
    DB 479 AGAGCCCTGATAGACCCCAATTTCTTCGGAGAGATCTGCTGACCACTGTGTCACCTC 538
    QY 324 ACTCCTTCAACAGATTTTTCATCTCTGCTGCTGCTCTTCAATTAACATAAACAAGTCCATCT 383
    DB 539 TCCAGTATCAAAAGGTTTTCATCTCTGCTGCTGCTTATTAATATCAACAGAGATTT 598
    QY 384 GGAAATGATTTTCAAGAGATTACTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 443

```



Db 599 GGAATAATGTTTGTGATGAGCTTCTTCCAGATTAGTGAAAGAGGAAGTGTGTTAATT 658  
 QY 444 ATGCGCAACTGCAGCTAATGACCTTTTCATTTATTCGAGTCATCTCTAGAGGATTCACCT 503  
 Db 659 ATGATCCAGTGCTCTTTTGTGACAGCATCTCTCTGAGGCGCTCTCCAAAGAAATTCACCT 718  
 QY 504 ATGGAAGTGTGTAGCTGAGTGGAATTCACGGATGCTCTCCCTCAAGACTACGAGCTTTAA 563  
 Db 719 ATGGTAAAGTTTGTGCGAGAGGCTAAGTTTCAAGAGTCTCTCTGAGCTTACATGCTCGGA 778  
 QY 564 TTCGAGCTTAAGTAAGAAAGAGTGTGAATTTGTTGACATTTACAGCGTTGAAGAGA 623  
 Db 779 TAATAGCACAGAGCTGCATCACTAATGACCTCCCTCACTATGAAACGGTGGAGCGTG 838  
 QY 624 CGGTGAGGAAGAGAGTGTGAAGAAGGCTGTGGTGTGGCGAGAGTGAATCTTAACA 683  
 Db 839 CTATTGACATAGGTGGAAGCTAAGCTAAGCTTGTGACAGAGTGAATTTAGGCG 898  
 QY 584 GTGATGACAAATGACAAT---GAAACCGTAATTTGATCCATGCTGCTCTAGCTTGT 740  
 Db 899 CTGAAGACACAGCGGCTCCACCAATGTACAAAGTAAGGCCCGAGTTTGGTGGCTGAACGT 958  
 QY 741 ACAAATTTGGTGTATCTCTCCACAGGAGGTTTCAGTTGAGTACTCTTGGCGCGTC 800  
 Db 959 ACAGCTACAGATCATGCCGTAAACAGAGGTTCAAGTAGCCTACTTCTGAGGAGAT 1018  
 QY 801 TAGACTGAAGCATTAACAATGCA 823  
 Db 1019 TGGATTGATTTTACGATTGTA 1041

RESULT 15  
 AK068983  
 LOCUS AK068983 1250 bp mRNA linear PLN 24-JUL-2003  
 DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J023001N13, full insert sequence.  
 ACCESSION AK068983  
 VERSION AK068983.1 GI:32979007  
 KEYWORDS FLI\_CDNA; CAP trapper.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE 1  
 AUTHORS The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Hayashida, K., Hayatsu, N., Hiramoto, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohtsuki, K., Shishiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nakikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Kusunegi, T., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Otsu, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
 Science 301 (5631), 376-379 (2003)

TITLE japonica rice  
 JOURNAL Science 301 (5631), 376-379 (2003)  
 MEDLINE 22752273  
 PUBMED 12869764  
 REFERENCE 2 (bases 1 to 1250)  
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Inamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

FEATURES  
 source  
 1..1250  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J023001N13"

ORIGIN  
 Query Match 25.2%; Score 257.4; DB 8; Length 1250;  
 Best Local Similarity 58.4%; Pred. No. 8.9e-50;  
 Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;  
 QY 24 TGGCCAAAGCAGCAACAAAGTCTGATTCCTGGGAATGTGTACACGCTAGCTTCTGTA 83  
 Db 245 TGACCAAGGAGGAGAAACAGAGATAGATCAAGTGAATACTGACCTTGGACACATTA 304  
 QY 84 GAGAGATTGTTGGTGGCAAGAGGATACCATCATTTATGGTCTCATTTAGAGAGCAAGT 143  
 Db 305 GAACCTCTCTGTTAGGCAAGAGAGACATCATATTCAGCTCTTTAGAGAGACACAGT 364  
 QY 144 TCCTTAGCAATTCACACCTATGATGAAGTATGCTCAATCCAGGTTTGTGGCT 203

TITLE Direct Submission  
 JOURNAL Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
 COMMENT This clone is one of the 28K full-length cDNA clones from japonica rice.  
 URL : http://cdna01.dna.affrc.go.jp/cDNA/  
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.  
 PAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusunegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nakikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
 and Genome Science Laboratory in Riken Genomic Sciences Center  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohta, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 1..1250

```

Db      365 TTGCTACAAATGCTGATATATATGATAAAATGCTTTCCATGTGGATGGATTTGATGGCT 424
Qy      204 CATTGTGGAATTTGTTGTAAGATPACAGAGGCCAATCAAGCTAAGCTGAGCTGGAAGATACA 263
Db      425 CTTTGGTTGAATTCATGTTTAGAGAAACCGAAATCTACATCAACAGGTTGGGAGATACA 484
Qy      264 AAAACCTGHAAGAAAGCGCTTCTCCAGAAAAATTTACCAACATCAATTTGCCATCTT 323
Db      485 AGAGCCCTGATGAGCACCATTCTTTCCGAGGATCTGCTGAACACCTGTTGCCACCTC 544
Qy      324 ACTCCTTCAAAACAGTTTTCATCTCTGGTGTCTCTCAATTAACATAAAACAAGTCCATCT 383
Db      545 TCCAGTATCCAAAGTTTTCATCTCTCTCAATTAACATAAAACAAGTCCATCT 604
Qy      384 GGAATATGTAATTCAAAGTACTTCCATTTGCTTCTCTCAATTTGCTGATCTAATTAATCAACAGGAGATTT 443
Db      605 GGAATATGTAATTTGATGAGCTTCTTCCAAAGATTTAGTGAAGAAAGGAGTGTGTTAAT 664
Qy      444 ATGCGCAACTGCAGCTTAATGACCTTTTCAATTTGCTGAGTCCATCTCTAGAAGGATTCCT 503
Db      665 ATGATCCAGTGTCTTTTGTGACAGATCTGCTTGGAGGGCTCTCCAAAGAAATTCCT 724
Qy      504 ATGGAAGTTTGTAGCTGAGTGGAATTCAGGATCTCTCAAGACTACGAGCCTTTAA 563
Db      725 ATGTRAAGTTTGTGCGCAGAGCTAAGTTTCAAGAGTCTCTGAAGCTTACATGCCTGCGA 784
Qy      564 TTCGAGCTAAGATAAAGAGGATTCATGAAATTTGACATTTACAGCGTTGAAGAGA 623
Db      785 TAATAGCACAGGACTGCGATCAACTAATGACCTCTCTCACTATGAACGGTGGAGCGTG 844
Qy      624 CGGTGAGGAAGAGAGTTGAAAGAGGCTGTGTTGTTGGCAGGAGTGAATCTTAACA 683
Db      845 CTATTGAACATAGGTGGAAGCTAAGCTAAGATCTTTGGACAGGAGGTGGATTTAGGCG 904
Qy      684 GTGATGCAATGACAAT---GAAACGGTAATTTGATCCATCAGTGGCTTCTAGCTTGT 740
Db      905 CTGAAGACACCGCGCTCCACCAATGTACAGATAAGGCCCAAGTTTGGTGGCTGAACCTGT 964
Qy      741 ACAAAATTTGGTGATACCTCTCAACAAGGAGTTCAAGTTGAGTACCTCTTTCGCCGCTC 800
Db      965 ACAGCTACAGGATCATGCGCTAACCAAGGAGTTCAAGTAGCCTACTTGTCTGAGGAGAT 1024
Qy      801 TAGACTGAGGCAATTACAATGCA 823
Db      1025 TGGATTGATTGTTTACGATTGTA 1047

```

Search completed: September 25, 2004, 00:33:08  
Job time : 4189 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 22:28:41 ; Search time 477 Seconds  
(without alignments)  
9084.197 Million cell updates/sec

Title: US-10-624-061-15  
Perfect score: 1020  
Sequence: 1 gcacgagagcttaattcaa.....atcttcaaaaaaaaaa 1020

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001as.\*  
5: Geneseq2001bs.\*  
6: Geneseq2002s.\*  
7: Geneseq2003as.\*  
8: Geneseq2003bs.\*  
9: Geneseq2003cs.\*  
10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	100.0	1020	7	ABX93058 Soybean c
2	485.6	45.6	525	7	ABX93053 Soybean c
3	327.2	32.1	1006	6	ABA99613 A. thalia
4	327.2	32.1	1006	6	ABA91400 Arabidops
5	327.2	32.1	1006	6	ABA91394 Arabidops
6	326.2	32.0	798	6	ABZ13167 Arabidops
7	326.2	32.0	993	6	ABA99614 Construct
8	257.4	25.2	1231	7	ABX93057 Rice chor
9	255	25.0	1274	3	AAC33324 Arabidops
10	254.8	25.0	1207	6	ABA91393 Arabidops
11	254.8	25.0	1207	6	ABA91399 Arabidops
12	251	24.6	1005	6	ABZ12932 Arabidops
13	239.4	23.5	1223	7	ABX93056 Corn chor
14	229.4	22.5	780	7	ABX93059 Wheat cho
15	211.4	20.7	1217	6	ABA91395 Arabidops
16	211.4	20.7	1217	6	ABA91401 Arabidops
17	206.2	20.2	579	7	ABX93054 Wheat cho
18	99.8	9.8	843	3	AAC81949 H. polymo
19	99.8	9.8	1655	3	AAC81950 H. polymo
20	98.8	9.7	5176	6	ABA91397 Arabidops
21	93.8	9.2	258	6	ABL71526 Corn tass
22	84	8.2	287	6	ABL72637 Corn tass
23	73.6	7.2	5099	6	AAD29643 Arabidops

## ALIGNMENTS

RESULT 1  
ABX93058  
ID ABX93058 standard; cDNA; 1020 BP.  
XX  
AC ABX93058;  
DT 22-MAY-2003 (first entry)  
XX  
DE Soybean chorismate mutase cDNA.  
XX  
KW Soybean; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;  
KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
XX  
OS Glycine max.  
XX  
FH Key Location/Qualifiers  
FT CDS 23..808  
FT /\*tag= a  
FT /product= "Chorismate mutase"  
XX  
PN US2002184658-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 03-DEC-1999; 99US-00454279.  
XX  
PR 04-DEC-1998; 98US-0110845P.  
XX  
PA (FALC/) FALCO S C.  
PA (FAMO/) FAMODU O O.  
PA (LEEC/) LEE J.  
XX  
DR Falco SC, Famodu OO, Lee J;  
DR WPI; 2003-328651/31.  
DR P-PSDB; ABU08094.  
XX  
PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
PT nucleic acids, useful for facilitating design and/or identifying  
PT inhibitors of those enzymes that may be used as herbicides and for  
PT producing antibodies.  
XX  
PS Claim 3; Page 22; 32pp; English.  
XX  
CC The invention discloses isolated polynucleotides encoding chorismate

Abx91396 Arabidops  
Abx93051 Corn chor  
Abx93052 Rice chor  
Aal50202 M sterili  
Aaf14469 Aspergill  
Aaf17891 Aspergill  
Abt19705 Aspergill  
Aah44222 Physcomit  
Ada71938 Rice gene  
Abx97892 DNA encod  
Ada71938 Rice gene  
Ab134175 Human inm  
Ab133676 Human inm  
Ab134576 Human met  
Ab170369 Chemical  
Ab132866 Human inm  
Aas63324 Chemical  
Ab133018 Human inm  
Ab133759 Human inm  
Ab132323 Human inm  
Aas61074 Human gen  
Continuation (9 of

mutase polypeptides. Aromatic amino acids must be included in the diets of animals. In the aromatic amino acid biosynthetic pathway chorismate is converted to anthranilate during tryptophan biosynthesis and is converted to prephenate, the branch point for tyrosine and phenylalanine biosynthesis. Chorismate mutase catalyses the conversion of chorismate to prephenate. Also disclosed are methods for selection of an isolated polynucleotide that affects the level of expression of a tyrosine biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid fragment encoding a tyrosine biosynthetic enzyme polypeptide and evaluating compounds for their ability to inhibit the activity of a tyrosine biosynthetic enzyme. The polypeptides can be used produce antibodies. Chorismate mutase and prephenate dehydrogenase are good targets for herbicides that will not affect animals, and overexpression of these enzymes may be used to increase the content of aromatic amino acid in food crops. The polypeptides may also be used to design and/or identify inhibitors of those enzymes that may be used as herbicides. The nucleic acids may be used to create transgenic plants, as probes for the genetic and physical mapping of the genes and as markers for traits linked to those genes. The sequence presented is a soybean chorismate mutase cDNA

Sequence 1020 BP; 322 A; 176 C; 223 G; 299 T; 0 U; 0 Other;

Query Match 100.0%; Score 1020; DB 7; Length 1020;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-268;  
 Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACGAGAGCTTAATTCAGAGATGCGCAAGAGCAGACAGCAAGCTCTGAGGAA 60  
 DB 1 GCACGAGAGCTTAATTCAGAGATGCGCAAGAGCAGACAGCAAGCTCTGAGGAA 60  
 QY 61 TGTGTACACGCTAGCTCTCTGAGAGAGGATTTGGTTAGCGAAGGATACCATCATTTA 120  
 DB 61 TGTGTACACGCTAGCTCTCTGAGAGAGGATTTGGTTAGCGAAGGATACCATCATTTA 120  
 QY 121 TGGTCTCATTCAGAGAGCCAGGTCCTAGCAATTCACACTAGTATGATGAAGATGTC 180  
 DB 121 TGGTCTCATTCAGAGAGCCAGGTCCTAGCAATTCACACTAGTATGATGAAGATGTC 180  
 QY 181 TCAATCCAGGGTTTTGTGGCTCAATTTGGTGGAAATTTGTTTAAAGATACAGAGCCAT 240  
 DB 181 TCAATCCAGGGTTTTGTGGCTCAATTTGGTGGAAATTTGTTTAAAGATACAGAGCCAT 240  
 QY 241 TCAAGCTAAGCTGGAAGATACAAAACCTCGAAGAAACGCTCTTCCAGAAATTT 300  
 DB 241 TCAAGCTAAGCTGGAAGATACAAAACCTCGAAGAAACGCTCTTCCAGAAATTT 300  
 QY 301 ACCACCATCAATTTGTCATCTTACTCTCTCAACAGATTTTTCATCTCTGGTCTTC 360  
 DB 301 ACCACCATCAATTTGTCATCTTACTCTCTCAACAGATTTTTCATCTCTGGTCTTC 360  
 QY 361 AATTAAACATAAACAAGTCCATCTGGAATATGTAATTTCAAGAGTTACTTCCATTCGTC 420  
 DB 361 AATTAAACATAAACAAGTCCATCTGGAATATGTAATTTCAAGAGTTACTTCCATTCGTC 420  
 QY 421 TACTTCGGGTGATGCGCACTATGCGCAATCTGAGCTTAATGACCTTTCATTATGCA 480  
 DB 421 TACTTCGGGTGATGCGCACTATGCGCAATCTGAGCTTAATGACCTTTCATTATGCA 480  
 QY 481 GTCCATCTCTAGAGGATTCATATGGAATTTGTTAGCTGAGGTGAATTCAGGATGC 540  
 DB 481 GTCCATCTCTAGAGGATTCATATGGAATTTGTTAGCTGAGGTGAATTCAGGATGC 540  
 QY 541 TCCTCAAGACTACGAGCTTTAATTCAGCTAAGGATTAAGAGGATTCATGAAATTTGT 600  
 DB 541 TCCTCAAGACTACGAGCTTTAATTCAGCTAAGGATTAAGAGGATTCATGAAATTTGT 600  
 QY 601 GACATTTACAGCGCTTGAAGAGAGCGGTGAGGAAGAGATTTGAAAAGAGCGCTGTGT 660  
 DB 601 GACATTTACAGCGCTTGAAGAGAGCGGTGAGGAAGAGATTTGAAAAGAGCGCTGTGT 660  
 QY 661 TGGCAGAGAGTGAATCTTAACAGTGATGCAATGCAATGCAATGCAATGCAATGCAATC 720  
 DB 661 TGGCAGAGAGTGAATCTTAACAGTGATGCAATGCAATGCAATGCAATGCAATGCAATC 720

661 TGGCAGAGAGTGAATCTTAACAGTGATGCAATGCAATGCAATGCAATGCAATGCAATC 720  
 721 ATCAGTGGCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCACGAGGTTTCAGGT 780  
 721 ATCAGTGGCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCACGAGGTTTCAGGT 780  
 781 TGACTACCTCTTGGCGGCTCTAGACTGAAGGCATTACAATGCGAGTTAGAAATTTAGAGAA 840  
 781 TGACTACCTCTTGGCGGCTCTAGACTGAAGGCATTACAATGCGAGTTAGAAATTTAGAGAA 840  
 841 TGGAGATGAATATGATGTTGTTTCAATGATTAAGCTCTTAAGTGATCCCTTTATTC 900  
 841 TGGAGATGAATATGATGTTGTTTCAATGATTAAGCTCTTAAGTGATCCCTTTATTC 900  
 901 CAATTCATCTAGCTGTTGATTCAGAAATATTTTGTAGCTATAGATTGTTACCTTTAT 960  
 901 CAATTCATCTAGCTGTTGATTCAGAAATATTTTGTAGCTATAGATTGTTACCTTTAT 960  
 961 TTTTCATCGGCTTTTATAGGAAAGGCATTATATATCATGATCTTCAAAAAAARAAA 1020  
 961 TTTTCATCGGCTTTTATAGGAAAGGCATTATATATCATGATCTTCAAAAAAARAAA 1020

RESULT 2  
 ABX93053  
 ID ABX93053 standard; cDNA; 525 BP.  
 XX  
 AC ABX93053;  
 XX  
 DT 22-MAY-2003 (first entry)  
 XX  
 DE Soybean chorismate mutase expressed sequence tag.  
 XX  
 KW Soybean; plant; gene; ss; chorismate mutase; expressed sequence tag; EST;  
 KW aromatic amino acid; diet; biosynthesis; chorismate; anthranilate;  
 KW tryptophan; prephenate; tyrosine; phenylalanine;  
 KW prephenate dehydrogenase; herbicide; food crop.  
 XX  
 OS Glycine max.  
 XX  
 Key Location/Qualifiers  
 PH 43.438  
 FT CDS /\*tag= a  
 FT /product= "Chorismate mutase"  
 FT /partial  
 FT /note= "No start or stop codon shown"  
 XX  
 XX US2002184658-A1.  
 XX  
 PD 05-DEC-2002.  
 XX  
 PF 03-DEC-1999; 99US-00454279.  
 XX  
 PR 04-DEC-1998; 98US-0110845P.  
 XX  
 PA (FALC/) FALCO S C.  
 PA (FAMO/) FAMODU O O.  
 PA (LEEJ/) LEE J.  
 XX  
 PI Falco SC, Famodu OO, Lee J;  
 XX  
 XX WPI; 2003-328651/31.  
 DR P-PSDB; ABU08089.  
 XX  
 PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
 PT nucleic acids, useful for facilitating design and/or identifying  
 PT inhibitors of those enzymes that may be used as herbicides and for  
 PT producing antibodies.  
 XX  
 PS Claim 3; Page 16; 32pp; English.  
 CC The invention discloses isolated polynucleotides encoding chorismate  
 CC mutase polypeptides. Aromatic amino acids must be included in the diets

of animals. In the aromatic amino acid biosynthetic pathway chorismate is converted to anthranilate during tryptophan biosynthesis and is converted to prephenate, the branch point for tyrosine and phenylalanine biosynthesists. Chorismate mutase catalyses the conversion of chorismate to prephenate. Also disclosed are methods for selecting an isolated polynucleotide that affects the level of expression of a tyrosine biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid fragment encoding a tyrosine biosynthetic enzyme polypeptide and evaluating compounds for their ability to inhibit the activity of a tyrosine biosynthetic enzyme. The polypeptides can be used produce antibodies. Chorismate mutase and prephenate dehydrogenase are good targets for herbicides that will not affect animals, and overexpression of these enzymes may be used to increase the content of aromatic amino acid in food crops. The polypeptides may also be used to design and/or identify inhibitors of those enzymes that may be used as herbicides. The nucleic acids may be used to create transgenic plants, as probes for the genetic and physical mapping of the genes and as markers for traits linked to those genes. The sequence presented is a soybean chorismate mutase expressed sequence tag (EST)

Db 249 TTCTCTCAGTGGTCTTTCCTGTCAGAGAGACAGAAATCATCCAGCTAAGGTAGGAGATA 308  
 Qy 262 CAAAAACCCCTGAGAAAAAGCCCTTCTCCGAGAAAAATTTACCACCATCAATTTGTGCCATC 321  
 Db 309 TGAATACCCGAGAGATCCCTTCTCTGTCAGAGACATTCCTCACTCGGTTTTCCTAC 368  
 Qy 322 TTTACTCTTCAACAGTTTTTGGATCTCTGGTCTGCTTCAATTAACATAAACAGTCCAT 381  
 Db 369 GCACAAATATCCATCGGCTTTGCACCTTAAGGCTCTATCTGTAAATTAACAAACAAAT 428  
 Qy 382 CTGCAAAATGATTTTCAAGAGTTACTTCCATTTGCTTGTCTTCCGCTGATGATGGCAA 441  
 Db 429 CTGGGATATTTACTTTAAGAAATGCTTCTTGTGTTGTCACACCTGGGATGATGGCAA 488  
 Qy 442 CTATGGCAAACTGACGATTAATGACCTTTCAATTTGCTGATCTCTAGAGGATTTCA 501  
 Db 489 CTATCCATCAACTGCTGCTAGTATGATCTCGCTGTTTACAGCTCTTTTCGAGAAGGATTTCA 548  
 Qy 502 CTATGGAAGTTGTAGCTGAGGTGAATTTTCAGGATGCTCTCAAGACTACGAGCTTT 561  
 Db 549 CTACGGTAATTTGTAGCTGAGTCAATTTTCAGATGCTTCCACAGATTACAGGCTTCG 608  
 Qy 562 AATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGTGACATTTACAGGCTTTGAAGA 621  
 Db 609 GATTCGGCTCAGATAGAGGCTTTGATGAAGCTGTTGACGTTTGAAGAAAGTAGAAGA 668  
 Qy 622 GACGCTGAGAGAGAGTTGAAAGAGGCTGTGGTGTGTTGGGAGGAGTGAATCTTAA 681  
 Db 669 AATGGTTAAGAGAGAGTGAGAAAGAGGAGAAACGTTTGGACAAAGAGTAAATTTCA 728  
 Qy 682 -----CAGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTCTAG 735  
 Db 729 CTCTGGCTATGGGATGAGAGTAAAGAGATATAAAGTGGATCCATTCCTTGCCTCTCG 788  
 Qy 736 CTGTGACAAAATGGGTGATACCTCTCACCAGGAGGTTTCAGTTGATGATCTCTTGG 795  
 Db 789 CATCTACGGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGATGATCTCTACG 848  
 Qy 796 CCGCTCTAGACTGAA 809  
 Db 849 TCGTCTCGATTGAA 862

RESULT 4

ABA91400/c  
 ID ABA91400 standard; DNA; 1006 BP.  
 AC ABA91400;  
 XX  
 DT 08-APR-2002 (first entry)  
 XX  
 DE Arabidopsis chorismate mutase-2 antisense polynucleotide.  
 XX  
 KW Chorismate mutase-2; enzyme; plant; antisense; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX WO200202798-A2.  
 XX  
 XX 10-JAN-2002.  
 XX  
 XX 22-JUN-2001; 2001WO-US020104.  
 XX  
 XX 05-JUL-2000; 2000US-00610040.  
 XX  
 XX (PARA-) PARADIGM GENETICS INC.  
 XX  
 XX Boyes DC, Davis KR, Woessner JP, Gorch J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 PI Kjemtrup S;  
 XX  
 XX WPI; 2002-154754/20.  
 XX

PT Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 XX targets which are essential for plant growth.  
 PS Claim 35; Page 71-72; 73pp; English.  
 XX  
 CC The present sequence is that of an antisense oligonucleotide that is  
 CC specific for Arabidopsis thaliana chorismate-mutase-2 (CM) mRNA. CM  
 CC catalyses the conversion of chorismate to prephenate in a branch of the  
 CC shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 CC is essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense and sense polynucleotides,  
 CC double-stranded RNA and ribozymes, and related expression cassettes and  
 CC vectors. The compositions are particularly useful for the modulation and  
 CC inhibition of plant growth, and for generating male sterile plants or  
 CC plants that produce seedless fruit  
 XX  
 SQ Sequence 1006 BP; 302 A; 226 C; 188 G; 290 T; 0 U; 0 Other;

Query Match 32.1%; Score 327.2; DB 6; Length 1006;  
 Best Local Similarity 54.2%; Pred. No. 7.4e-79;  
 Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
 Qy 22 AATGGCCAAAGCAGACAGAAACAAAGTCTGATTTCTGGGAATGTACACGCTAGCTTCTGT 81  
 Db 938 AAGAGTCTTCGAATCGGATTCGGGTTCTGGTCTTCCAAATGTACTGAGTCTTACATAA 879  
 Qy 82 GAGAGAGATTTGGTAGGCAAGAGATACCATCATTTATGCTCTCATTTGAGAGAGCCAA 141  
 Db 878 CAGAGAAATCTGTTAGTAGGCAAGAGACACCATCGCTTTCAGCTTGTATGAGAGAGCTAA 819  
 Qy 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAAAAGTATGCTCAATCCAGGGTTTTGTGG 201  
 Db 818 GTTTCGACTCAATCTCTCTGTTTCGAGGATCTCGTGTCTAGATTCTGGAAGTTCTC 759  
 Qy 202 CTCATTCGTTGGAATTTGTTTAAAGAAATACAGAGGCGCATTCAGAGCTAAGCTGGAAGATA 261  
 Db 758 TTCTCTCACTGAGTTTTTCTGTCAGAGAGACAGAAATCATCCAAAGCTAAGGTAGGAGATA 699  
 Qy 262 CAAAAACCCCTGAGAAACCGCTTTCTCCAGAAAAATTTACCACCATCAATTTGCGCATC 321  
 Db 698 TGAATACCCGAAAGAGATCCCTTTCTCTTGAGACATTCCTCACTCGGTTTTCTTAC 639  
 Qy 322 TTACTCTTCAACAGTTTTTGGCATCTGGTGTCTGCTTCAATTAACATAAACAGTCCAT 381  
 Db 638 GCACAAATATCCATCGGCTTTGCAACCTAAGGCTCTATCTGTTAAACATTAACAAACAAAT 579  
 Qy 382 CTGGAATATGTTTCAAGAGTACTTCCATTTGCTTGTCTTCTCGGTTGATGCGCAA 441  
 Db 578 CTGGATATTTACTTTAAAGAAATGCTTCTTCTTTGTTGCAACCTGGGATGATGGCAA 519  
 Qy 442 CTATGCGCAAACTGACGCTAATGACCTTTTCATTTATTCAGTCCATCTCTAGAGAGTTCA 501  
 Db 518 CTATCCATCAACTGCTGTAGTGTATCTGGCTGTTTACAGCTCTTTCGAGAAGGATTTCA 459  
 Qy 502 CTATGGAAGTTTGTAGCTAGGTGAATTTACAGGATGCTCTCCACAGCTTACCGAGCTTT 561  
 Db 458 CTACGGTAAATTTGTAGCTAGGTCAAAATTCAGAGATGCTTCCACAGATTACGAGGCTGC 399  
 Qy 562 AATTTCGAGCTAAGGATAAAGAGGATTTGATGAAGCTTTGACCTTTTGAGAAAGTAGAAGA 621  
 Db 398 GATTTCGGCTCAGGATAGAGAGGCTTTGATGAAGCTTTGACCTTTTGAGAAAGTAGAAGA 339  
 Qy 622 GACGCTGAGGAGAGAGTTGAAAAGAGGCTGTGGTGTGTTGGCAGAGAGTGAATCTTAA 681  
 Db 338 AATGGTTAAGAGAGAGTGCAGAAAGAGAGAAACGTTTGGACAAAGAGTAAAAATTTCAA 279  
 Qy 682 -----CAGTGATGACAAATGACAAATGAAACCGTAAATTTGATCCATCAGTGGCTCTAG 735  
 Db 278 CTCTGGCTATGGCGATGAGAGTAAAGAGAGTATAAAGTGGATCCATTTGCTTCCCTCTCG 219

QY	736	CTTGTACAAAATTTGGTGTAGTACCTCTCACCAAGGAGGTTTCAGGTTGAGTACCTCTTGGG	795
DB	218	CATCTACGGGGATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTTCTACG	159
QY	796	CCGCTCTAGACTGAA	809
DB	158	TCGTCCTCGATTGAA	145
RESULT 5			
ABA91394			
ID	ABA91394	standard; cDNA; 1006 BP.	
XX	AC	ABA91394;	
DT	08-APR-2002	(first entry)	
DE	Arabidopsis chorismate mutase-2 cDNA.		
XX	KW	Chorismate mutase-2; herbicide; plant; enzyme; gene; ss.	
OS	Arabidopsis thaliana.		
PN	WO200202798-A2.		
PD	10-JAN-2002.		
PF	22-JUN-2001; 2001WO-US020104.		
PR	05-JUL-2000; 2000US-00610040.		
XX	(PARA-) PARADIGM GENETICS INC.		
PI	Boyes DC, Davis KR, Woessner JP, Goriach J, Hamilton CM;		
PI	Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;		
PI	Kjemtrup S;		
XX	WPI; 2002-154754/20.		
DR	Identifying compounds useful for modulating and inhibiting plant growth		
PT	and development, by using chorismate synthase and chorismate mutase as		
PT	targets which are essential for plant growth.		
XX	Claim 37; Page 64-65; 73pp; English.		
XX	The present sequence is that of a full-length cDNA encoding Arabidopsis thaliana chorismate mutase-2 (CM), an enzyme that catalyses the conversion of chorismate to prephenate in a branch of the shikimate pathway leading to phenylalanine and tyrosine biosynthesis. CM is essential for plant growth; inhibition of CM gene expression in plant seedlings results in severe chlorosis, reduced growth and developmental abnormalities. The invention provides compositions for the modulation of plant growth or development using CM antisense (see ABA91400) and sense polynucleotides, double-stranded RNA and ribozymes, and related expression cassettes and vectors. The compositions are particularly useful for the modulation and inhibition of plant growth, and for generating male sterile plants or plants that produce seedless fruit. CM is also a target for the identification of herbicides, and methods are provided for using CM in the identification of herbicide candidates. Methods are also provided for identifying compounds that stimulate the expression or function of CM. These can be used to promote growth and development		
XX	Sequence 1006 BP; 290 A; 188 C; 226 G; 302 T; 0 U; 0 Other;		
SQ	Query Match 32.1%; Score 327.2; DB 6; Length 1006; Best Local Similarity 64.2%; Pred. No. 7.4e-79; Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;		
QY	22	AATGCCAAGCAGCAGACAAAGTCCTGATTCCTGGGAATGCTACACGGCTAGCTTCTGT	81
DB	69	AAGAGTCTTCGAATCGGATTCGGGTCTGGTGTCTCCAAATGACTGACTTACTTAAT	128

QY	82	GAGAGGAGATTGGTTAGGCAAGAGAGATACCATCATTTATGGTCTCATTTGAGAGCCAA	141
DB	129	CAGAGATCGTTGATTAGGCAAGAGAGACACCATCGTCTTACGTTGATCGAGAGACTAA	188
QY	142	GTTCCTAGCAATCTCACCTATGATGAAAGATATGCTCAAATCCAGGGTTTTTGTGG	201
DB	189	GTTCCTAGCAATCTCTGCTTTCGAGGAATCTCGTTGCTAGATTTCTGGAAGTTTCTC	248
QY	202	CTCATTTGGTGGAAATTTGTTTAAAGATACAGAGCCATTCAAGCTAAGGCTGGAGATA	261
DB	249	TTCTCTCACTAGTTCCTGTCAGAGACAGAAATCATCAAGCTAAGTAGTAGAGATA	308
QY	262	CAAAAACCTCGAAGAAAGCCCTTCTTCCAGAAATTTACCAATCAATTTGTGCCATC	321
DB	309	TGAATACCCCGAAGAGAAATCTTCTTCTTCTGAGAATCTCTCACTCGGTTTTTCTAC	368
QY	322	TTACTCTCTCAACACAGTTTTCGCTCTGCTGCTCTCAATTAACATAAACAAGTCCAT	381
DB	369	GCACAAATATCCATCGGCTTTGCACCCCTAAGCTCTATCTGTTAAATTAACAAACAAT	428
QY	382	CTGGAAATGTATTTCAAAGAGTTACTTCCATTTGCTTGTCTACTTCGGGTGATGAGCAA	441
DB	429	CTGGGATATTTACTTTAAAGAAATTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	488
QY	442	CTATGGGCAACTGCAGCTAATGACCTTTCATTTGCTGCTCTCTAGAGGATTTCA	501
DB	489	CTATCCATCACTGCTGCTAGTATCTGCTCTTCAAGCTCTTTCGAGAGGATTTCA	548
QY	502	CTATGGAAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCTCAAGACTACGAGCCTTT	561
DB	549	CTACGGTAATTTGTAGCTGAGTCAATTCAGAGATGCTCCACAAGATTACGAGCCTGC	608
QY	562	AATTCAGTAAGGATAAAGAGGATTTGATGAAATTTGTTGATTTACATTTACAGCGTTGAGA	621
DB	609	GATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGAAGA	668
QY	622	GACGGTGAAGAGAGAGTTGAAAGAGGCTCTGCTGTTTGGGCGAGGAGTGAATCTTAA	681
DB	669	AATGTTAAGAGAGAGTCCAGAGAGAGAGGTTTGGACAAGAGTAAATTTCAA	728
QY	682	-----CACTGATGACAAATGACAATCAAAACCGTAAATTTGATTCATCAGTGGCTTAG	735
DB	729	CTCTGGCTATGCGATGAGAGTAAGAGAAAGTATAAAGTGAATCCATTGCTTGCCTCTCG	788
QY	736	CTTGTAACAAATTTGGTGTATACCTCTCACCAGGAGGTTTCAGGTTGAGTACCTCTTGG	795
DB	789	CATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTTCTTACG	848
QY	796	CCGCTCTAGACTGAA	809
DB	849	TCGTCCTCGATTGAA	862
RESULT 6			
ABZ13167			
ID	ABZ13167	standard; DNA; 798 BP.	
XX	AC	ABZ13167;	
XX	AC	ABZ13167;	
DT	21-JAN-2003	(first entry)	
DE	Arabidopsis thaliana stress regulated gene SEQ ID NO 972.		
XX	Arabidopsis thaliana; plant; gene; stress; transgenic; ds.		
OS	Arabidopsis thaliana.		
XX	PN	WO200216655-A2.	
XX	PD	28-FEB-2002.	
XX	PF	24-AUG-2001; 2001WO-US026685.	
XX	XX		









Db 239 TGACCAAGGAGGAGAAACAGAGGATAGATCAAAAGTGAAATACTGACCTTGGACAACATTA 298  
Qy 84 GAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGTTCTCATTTGAGAGAGCAAGT 143  
Db 299 GAACCTCTTGGTTAGGCAAGAGAGAGATCATATTCAGCTTCTTAGAGAGAGACAGT 358  
Qy 144 TCCTAGCAATTTCTCACCTATGATGAAGAAAGTATGCTCAAAATCCAGGGTTTTTGCGT 203  
Db 359 TTTGCTACAATGCTGATATATATGATAAAATGCTTTCCATGCGATGATTTGATGGCT 418  
Qy 204 CATTTGGTGAATTTGTTTGAAGATACAGAGGCCATTCAGCTAAGGCTGGAAGATACA 263  
Db 419 CTTTGGTGAATTCATGGTTAGAGAACCGAAACTATACATCAACAGGTTGGAGATACA 478  
Qy 264 AAAACCTCTGAAGAAACGCTTCTTCCAGAAAATTTACCAATCAATTTGCGCATCTT 323  
Db 479 AGAGCCCTGATGAGCACCCATCTTTCGAGGATCTGCTGAACCACTGTTGCCACCTC 538  
Qy 324 ATCTCTTCAACAGTTTTCATCTCTGCTGCTGCTTCAATTAACATAAACAAGTCAATCT 383  
Db 539 TCCAGTATCCAAAGGTTTTCATCTCTGCTGATTTCTATTAATATCAACAAGGAGATT 598  
Qy 384 GGAATATGATTTCAAGAGTTACTTCCATTTGCTTCTTCTGCTGCTGATGATGCACT 443  
Db 599 GGAATATGATTTTGTATGAGCTTCTTCCAGATTAAGTGAAGAGGAGTGAATTT 658  
Qy 444 ATGCGAAATGAGCTGATGATGATGCTTCTTCAATTTATGAGTCCATCTCTGAAGAGATCT 503  
Db 659 ATGATCCAGTCTCTTGTGACACGATCTGCTTGCAGGCGCTCTCCAAAGAAATTCAT 718  
Qy 504 ATGGAAGTTTGTAGCTGAGTCAATTCAGGATGCTCTCAAGACTACGAGCTTTTAA 563  
Db 719 ATGTAAGTTTGTGAGAGGCTTAAGTTTCAAGAGTCTCTGAGCTTACATGCTTGGCA 778  
Qy 564 TTCGAGCTAAGGATAAGAGGATTTGATGAATTTGTGACATTTAAGCGTTGAAGAGA 623  
Db 779 TAATAGCACAGGATCGGATCAACTAATGCACTCTCACTATGAACGGTGGAGCGTG 838  
Qy 624 CGTGAGGAGAGAGTTGAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 683  
Db 839 CTATTGAACATAGGTTGAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCTTAAGCT 898  
Qy 684 GTGATGACATGACAAAT---GAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGT 740  
Db 899 CTGAGACACAGCGGCTCCACCAATGTACAAGATAGGCCAGTTTGGTGGCTGAACGT 958  
Qy 741 ACAAAATTTGGTGTATACCTCTCCACAGGAGTTTCAAGTGAAGTACCTTTCGCGCGTC 800  
Db 959 ACAGCTACAGGATCATGCGCTTAACAGAGGTTTCAAGTGAAGTGAAGTGAAGTGAAG 1018  
Qy 801 TAGACTGAAGGCTTCAATGCA 823  
Db 1019 TGGATTGATTGTTAGATTGTA 1041

## RESULT 9

AAC33324

ID AAC33324 standard; DNA; 1274 BP.

XX XX

AC AAC33324;

XX XX

DT 17-OCT-2000 (first entry)

XX XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 2614.

XX XX

KW Hybridisation assay; genetic mapping; gene expression control;

XX XX

KW protein identification; signal transduction pathway; metabolic pathway;

XX XX

KW promoter; termination sequence; ss.

XX XX

OS Arabidopsis thaliana.

XX XX

PN EP1033405-A2.

XX PD 06-SEP-2000.  
XX XX 25-FEB-2000; 2000EP-00301439.  
XX XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 03-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137328P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139452P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140931P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.



Db 747 TCAAAGAGATCCATTACGGTAAATTTGTTGAGAGCTAAATTTCAAGCCCTCACCCGAA 806  
QY 548 GACTACAGCCTTAAATTCAGCTAAGGATTAAGAGAGGATTTGATGAATTTGACATTT 607  
Db 807 GCATACGAGTCCGCCATCAAGACCAAGATAGAGGATGCACTGATGATATGCTGATTC 866  
QY 608 ACAAGCGTTGAAGAGACGGTGAAGAGAGAGTTGAAAGAGAGCTGTGTTGTTGGGAG 667  
Db 867 CCGACTGTGGAAGATGGGATTAAGAGAGAGTTGATGAAACCGGACATACGGGCAA 926  
QY 668 GAAGTGAATCTTAACAGTGAATCAATGAAACCGTAAATTTGATCCATC --- 723  
Db 927 GAAGTGAAGTTCGGATGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 986  
QY 724 -----AGTGGCTTCTAGCTTTGATCAAAATTTGGGTGATACCTCTCA 764  
Db 987 GTTTACAAATCATGCTGATCTTAGTTGTGACTTATATGAGATTGATGCTGCTTTAA 1046  
QY 765 CCAAGGAGGTTTCAAGTTGAGTACCTCTTGGCCGCTCTAGACTGAAGGCAATTA 819  
Db 1047 CAAAAGAGTTCAAGTGGAGTACTTGTCTAGAGAGACTGGACTGAGGCAACAA 1101

RESULT 10  
ID ABA91393 standard; cDNA; 1207 BP.

AC ABA91393;  
XX  
DT 08-APR-2002 (first entry)  
XX Arabidopsis chorismate mutase-1 cDNA.  
DE Arabidopsis chorismate mutase-1; herbicide; plant; enzyme; gene; ss.  
XX Chorismate mutase-1; herbicide; plant; enzyme; gene; ss.  
XX Arabidopsis thaliana.

OS Arabidopsis thaliana.  
XX WO200202798-A2.  
XX 10-JAN-2002.  
XX 22-JUN-2001; 2001WO-US020104.  
XX 05-JUL-2000; 2000US-00610040.  
XX (PARA-) PARADIGM GENETICS INC.  
XX Boyes DC, Davis KR, Woessner JP, Gorlach J, Hamilton CM;  
PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
PI Kjemtrup S;  
XX WPI; 2002-154754/20.  
XX  
XX Identifying compounds useful for modulating and inhibiting plant growth  
XX and development, by using chorismate synthase and chorismate mutase as  
XX targets which are essential for plant growth.  
XX  
XX Claim 37; Page 64; 73pp; English.

XX The present sequence is that of a full-length cDNA encoding Arabidopsis  
XX thaliana chorismate mutase-1 (CM), an enzyme that catalyses the  
XX conversion of chorismate to prephenate in a branch of the shikimate  
XX pathway leading to phenylalanine and tyrosine biosynthesis. CM is  
XX essential for plant growth; inhibition of CM gene expression in plant  
XX seedlings results in severe chlorosis, reduced growth and developmental  
XX abnormalities. The invention provides compositions for the modulation of  
XX plant growth or development using CM antisense (see ABA91399) and sense  
XX polynucleotides, double-stranded RNA and ribozymes, and related  
XX expression cassettes and vectors. The compositions are particularly  
XX useful for the modulation and inhibition of plant growth, and for  
XX generating male sterile plants or plants that produce seedless fruit. CM  
XX is also a target for the identification of herbicides, and methods are

CC provided for using CM in the identification of herbicide candidates.  
CC Methods are also provided for identifying compounds that stimulate the  
CC expression or function of CM. These can be used to promote growth and  
CC development.  
XX  
SQ Sequence 1207 BP; 343 A; 232 C; 278 G; 354 T; 0 U; 0 Other;  
Query Match 25.0%; Score 254.8; DB 6; Length 1207;  
Best Local Similarity 57.4%; Pred. No. 4.6e-59;  
Matches 494; Conservative 0; Mismatches 342; Indels 24; Gaps 1;  
QY 68 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGCCAGAGGATACCATTTATGTGCTC 127  
Db 253 ACTCTTGAAGGATATTAGAACTCTTTGATCCGTCAGAGGACAGCATTTATTTGGCTA 312  
QY 128 ATTGAGAGAGCCAGTTCCCTAGCAATTTTCACACCTTATGATGAAAGTATGCTCAATC 187  
Db 313 TTGAGAGAGCCAGTACTGTTCATATGCTGATCTATGATCTCTGCTTTGACATG 372  
QY 188 CAGGTTTTTGTGCTCATTTGGTGGAAATTTGTTTGAAGATACAGAGGCCATTCAGCT 247  
Db 373 GATGTTTCAATGCTTCTTTGGTTGATCATGTTTAAAGGCACTGAGAGCTTCAGCT 432  
QY 248 AAGCTCGAAGATACAAAACCCCTGAGAAAAGCCCTTTTCCAGAAAATTTTACCACCA 307  
Db 433 AAGTTGTTAGTTTATGAGTCTCTGATGACATCTTTCTTCCCTGATGATCTACAGAG 492  
QY 308 TCAATTTGTCATCTTACTCTTCAACAGTTTTCGATCTCTGCTGCTGCTTCAATTAAC 367  
Db 493 CCTATTTGCTCTCTTCTTCACTAGTACCCAAAGGTTTGCATTTTCTGCTGATTCGTAAC 552  
QY 368 ATAAACAGTCCATCTCTGAAATGTAATTTCAAGAGTTTACTTCCATTTGCTTACTTCG 427  
Db 553 ATAAACAGAGATATGGAATGTAATTTCAAGAGTTTACTTCCATTTGCTTACTTCG 612  
QY 428 GGTGATGATGCACTATGCGCAAACTGCACTTAAGAACCTTTTCAATTTTGAAGTCCATC 487  
Db 613 GGCATGATGTTACTTACCGCTCAACAGCTGTCTGTGACGCTATCTGCTTCACTGCTC 672  
QY 488 TCTAGAGGATTCATATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCCTCAA 547  
Db 673 TCAAAGAGATTCATACCGTAAATTTGTTGAGAGCTAAATTTCAAGCTCACCGAA 732  
QY 548 GACTACGAGCTTTAAATTCGAGCTAAGGATAAAGAGAGGATTTGATGAAATTTGTCATTT 607  
Db 733 GCATACGAGTCCGCCATCAAGCAACAGATGCGACTGATGGATATGCTGACATTC 792  
QY 608 ACAAAGCTTTGAGAGAGCGTGAAGAGAGAGTTGAAAGAGAGGCTGTGTTGTTGGGAG 667  
Db 793 CCGACTGTGGAAGATGCGATAAAGAGAGAGTTGAGATGAAACCCCGAACATACGGGCAA 852  
QY 668 GAAGTGAATCTTAAACAGTGTATGCAATGACAATGAAATGAAATGAAATGAAATGAAAT 705  
Db 853 GAAGTGAATCTTAAACAGTGTATGCAATGACAATGAAATGAAATGAAATGAAATGAAAT 912  
QY 706 --CCGTAAATTTGATCCATCAGTGGCTTCTAGCTTTGATGAAATTTGGGTGATACCTTC 763  
Db 913 GTTTACAAATCAGTCCGATCTTAGTTGGTGAATTTATATGAGATTGATGATGATGATGAT 972  
QY 764 ACCAAGAGGTTTCAAGTTGATGATCTCTTGGCCGCTCTAGACTGAAGGCAATTAACATGCA 823  
Db 973 ACAAAGAGGTTTCAAGTTGATGATCTCTTGGCCGCTCTAGACTGAAGGCAATTAACATGCA 1032  
QY 824 GTTGAATTTAAGAAATGGAAGATGAATGATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTT 893  
Db 1033 AACAAATGCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1092  
QY 884 AAGTGAATCTTAAATGCA 903  
Db 1093 TTACTTACAAATATTTCTAA 1112

## ABA91399/c

ID ABA91399 standard; DNA; 1207 BP.

XX AC ABA91399;

XX DT 08-APR-2002 (first entry)

XX DE Arabidopsis chorismate mutase-1 antisense polynucleotide.

XX DX Chorismate mutase-1; enzyme; plant; antisense; ss.

XX XX Arabidopsis thaliana.

XX OS WO200202798-A2.

XX PN 10-JAN-2002.

XX PD 22-JUN-2001; 2001WO-US020104.

XX PF 05-JUL-2000; 2000US-00610040.

XX PR (PARA-) PARADIGM GENETICS INC.

XX PA Boyes DC, Davis KR, Woessner JP, Goriach J, Hamilton CM;

XX PI Hoffmann NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;

XX PI KJ entrap S;

XX DR WPI; 2002-154754/20.

XX PT Identifying compounds useful for modulating and inhibiting plant growth

XX PT and development, by using chorismate synthase and chorismate mutase as

XX PT targets which are essential for plant growth.

XX PS Claim 35; Page 71; 73pp; English.

XX The present sequence is that of an antisense oligonucleotide that is  
 CC specific for Arabidopsis thaliana chorismate-mutase-1 (CM) mRNA. CM  
 CC catalyses the conversion of chorismate to prephenate in a branch of the  
 CC shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 CC is essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense and sense polynucleotides,  
 CC double-stranded RNA and ribozymes, and related expression cassettes and  
 CC vectors. The compositions are particularly useful for the modulation and  
 CC inhibition of plant growth, and for generating male sterile plants or  
 CC plants that produce seedless fruit

XX SQ Sequence 1207 BP; 354 A; 278 C; 232 G; 343 T; 0 U; 0 Other;

Query Match 25.0%; Score 254.8; DB 6; Length 1207;

Best Local Similarity 57.4%; Pred. No. 4.6e-59;

Matches 494; Conservative 0; Mismatches 342; Indels 24; Gaps 1;

QY 68 ACGGTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGGATACCATATTATGGTCTC 127

DB 955 ACTCTTGAAGGTATTAGAACTCTTTCATCGTCAAGAGGACACATTATATTGGGCTA 896

QY 128 ATTGAGAGAGCCAGTTCCCTAGCAATCTCAGACCTATGATGAAAAGTATGCTCAATC 187

DB 895 TTGAGAGAGAGCCAGTACTGTGTAAGTCTGATATCTATGATCTCTGCTTTTGACATG 836

QY 188 CAGGCTTTTGTGGCTCATTTGGTGGAAATTTTGTGTAAGATACAGAGGCCATTCAGCT 247

DB 835 GATGGTTCAATGGTCTTTGGTTGAGTACATGGTTAAAGGCACTGAGAAGCTTCAGCT 776

QY 248 AAGCTGGAAGATACAAAACCTTGAGAGAAACCGCTTCTTCCAGAAAATTTACACCA 307

DB 775 AAGTTGGTAGGTTAAGAGTCCCTGATGAACATCCTTCTTCCCTGATGATCTACAGAG 716

QY 308 TCAATTGGCCATCTTACTCTCTCAACAGATTTTGTGATCTCTGCTGCTCAATTAC 367

DB 715 CCTATGTCCTCTCTCTTCAGTACCCAAAGGTGTGCAATTTGCTGCTGATTCGATNAAC 656

QY 368 ATAAACAAGTCCATCTCGAAATATGTTATTCBAAGAGTTACTTCCATTGCTTCTACTTCG 427  
 DB |||||  
 QY 428 GGTGATGATGGCAACTATGCGAAACTGCGAAGCTTAATGACCTTTTCATTTATTCAGTCCATC 487  
 DB |||||  
 QY 595 GCGGATGATGTTAACTACCGCTCAACAGCTGCTGTGACGCTATCTGCGCTTCAGTGTCTC 536  
 DB |||||  
 QY 488 TCTAGAAGGATTCACCTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCTCAA 547  
 DB |||||  
 QY 535 TCAAGAGAAATCCATTACCGTAAATTTGTCAGAGAGCTAAATTTCAAGCCCTCACCCGAA 476  
 DB |||||  
 QY 548 GACTACGAGCCTTTAATTCGAGCTAAGGATAAAGAGGATTCATGAAATTTGTCATTT 607  
 DB |||||  
 QY 475 GCATACGAGTCCGCCATCAAGCACAAGATAAGGATCGACTGATGATATGCTGACATTC 416  
 DB |||||  
 QY 608 ACAAGCGTTGAAGAGACGGTGAAGAGAGTTGAAAGAGAGCTGTGCTGTTTGGGCGAG 667  
 DB |||||  
 QY 415 CCGACTGTGGAAGATGCGATAAAGAGAGAGTTGAGATGAAAACCCGAAACATACGGGCAA 356  
 DB |||||  
 QY 668 GAAGTGAATCTTAAACAGTGTGATGACATGCAATGCAATGAAAA----- 705  
 DB |||||  
 QY 355 GAAGTGAAGTTGGGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 296  
 DB |||||  
 QY 706 --CCGTAATTTGATCCATCAGTGGCTTCTAGCTGTGACAAAATTTGGGTGATACCTCTC 763  
 DB |||||  
 QY 295 GTTTACAAAATCAGTCCGATCTTAGTTGGTGACTTATATGAGAGATTGGATCATGCTTTTA 236  
 DB |||||  
 QY 764 ACCAAGAGGTTTCAGGTTGAGTACCTCTTGGCGCTCTAGACTGAGGCAATTACATGCA 823  
 DB |||||  
 QY 235 ACAAAGAGGTTCAAGTGGGAGTACTTGTCTCAGAGAGCTGGAGCTGAGGCAACAAATA 176  
 DB |||||  
 QY 824 GTTAGAATTTAGAAGAAATGGAAGATGAATATGATTTGTTTCAATGATTAAGCTCTT 883  
 DB |||||  
 QY 175 AACAAATATGCTTTGGTAGTACAGTAGAAGGTTTTTGAATGCTCTTTGGTTTTTTTTT 116  
 DB |||||  
 QY 884 AAGTGATCTTTATTGCCAA 903  
 DB |||||  
 QY 115 TTACTTTACAATATTTCTAA 96  
 DB |||||

## RESULT 12

ABZ12932

ID ABZ12932 standard; DNA; 1005 BP.

XX AC ABZ12932;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 737.

XX KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX OS Arabidopsis thaliana.

XX PN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US026685.

XX PR 24-AUG-2000; 2000US-0227866P.

XX PR 26-JAN-2001; 2001US-0264647P.

XX PR 22-JUN-2001; 2001US-0300111P.

XX XX (SCRI) SCRIPPS RES INST.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Harper JF, Kreps J, Wang X, Zhu T;

XX DR WPI; 2002-304127/34.

PT Identifying a stress condition to which a plant cell has been exposed and  
PT producing plants with increased tolerance to these abiotic stresses.  
PS Claim 144; SEQ ID NO 737; 577bp + Sequence Listing; English.  
XX  
XX The invention relates to identifying a stress condition to which a plant  
XX cell has been exposed, comprising: (a) contacting nucleic acid  
XX representative of expressed polynucleotides in the plant cell with an  
XX array or probes representative of the plant cell genome; and (b)  
XX detecting a profile of expressed polynucleotides in the plant cell  
XX characteristic of a stress response. The method is useful in the  
XX production of transgenic plants, cells and seeds and in producing plants  
XX with increased tolerance to abiotic stress. The present sequence is that  
XX of an Arabidopsis thaliana stress regulated gene (AB212196-AB212197) used  
XX in methods of the invention. Note: The sequence data for this patent is  
XX not represented in the printed specification but is based on sequence  
XX information supplied to Derwent by the European Patent Office  
SQ Sequence 1005 BP; 287 A; 205 C; 235 G; 278 T; 0 U; 0 Other;  
Query Match 24.6%; Score 251; DB 6; Length 1005;  
Best Local Similarity 59.6%; Pred. No. 4.7e-58;  
Matches 456; Conservative 0; Mismatches 285; Indels 24; Gaps 1;  
Qy 68 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGCGAAGAGATACCATCATTTATGCTCTC 127  
Db |||||  
Qy 241 ACTTTGAAGATTAGAACTCTTTGATCCGTCAGAGACAGCATTTATTTGGCTA 300  
Db |||||  
Qy 128 ATTGAGAGAGCAAGTTCCTAGCAATTCACACCTATGATGAAAGTATGCTCAATC 187  
Db |||||  
Qy 301 TTGAGAGAGAGCAAGTACTGTTCAATGCTGATATGATGCTCTCTTTGACATG 360  
Db |||||  
Qy 188 CAGGTTTTTGTGCTCATTTGGTGAATTTGTTAAGAAATACAGAGCCATTTCAAGCT 247  
Db |||||  
Qy 361 GATGTTTCAATGGTCTTTGTTGAGTACATGTTTAAAGGCACTGAGAGCTTCAAGCT 420  
Db |||||  
Qy 248 AAGCTGGAAGATACAAAAACCTGAGAAAAAGCGCTTCTCCGAAAAATTTACACCA 307  
Db |||||  
Qy 421 AAGTTGGTAGTTTAAAGATGCTCTGATGAACATCTTCTCCCTGATGATCTACAGAG 480  
Db |||||  
Qy 308 TCATTTGTCATCTTACTCTTCAACAGTCTTTTGCATCTCTGCTCTCAATTAAC 367  
Db |||||  
Qy 481 CCTATGTTGCTCTCTTCTAGTACCCAAAGGTTGCAATTTGCTGCTGATGATAAC 540  
Db |||||  
Qy 368 ATAAACAGTCCATCTGGAATAATGTTTCAAGAGTTACTTCCATGTTGCTACTTCG 427  
Db |||||  
Qy 541 ATAAACAGAGATATGGAACATGTACTTCAGAGACCTTGTTCGAAGACTTGTGAAGAA 600  
Db |||||  
Qy 428 GGTGATGATGCAACTATGCGCAACTGCACTTAATGACCTTCAATTATGAGTCCATC 487  
Db |||||  
Qy 601 GCGATGATGTTACTACGCTCAACAGCTGTCTGTGACGCTATCTGCTTCACTGCTC 660  
Db |||||  
Qy 488 TCTAGAAGATTCACTATGAAAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAA 547  
Db |||||  
Qy 661 TCAAGAGATCCATTACGCTAAATTTGTTGCAAGAGCTAAATTTCAAGCTCCACCGAA 720  
Db |||||  
Qy 548 GACTACAGCTTTAATTCAGCTAGGATTAAGAGAGGATGATGAATTTGACATTT 607  
Db |||||  
Qy 721 GCATACAGCTCCGCTCAACAGCAAGATAGGATGCACTGATGATATGCTGACATTC 780  
Db |||||  
Qy 608 ACAAGCTTGAAGAGAGCGTGAAGAGAGATTTGAAAGAGGCTGTGTTTGGGCGAG 667  
Db |||||  
Qy 781 CCGACTGTGAAGATGCAATGAAGAGAGATTTGAGATGAACCCGACATACGGCA 840  
Db |||||  
Qy 668 GAAGTGAATTTAACTGATGATCAATGCAATGAATAA-----705  
Db |||||  
Qy 841 GAAGTGAAGTTGGATGAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 900  
Db |||||  
Qy 706 --CGTAAATTTGATCCATCAGTGGCTTCTAGCTTGTACAAATTTGGGATACCTCTC 763  
Db |||||  
Qy 901 GTTTACAAATCAGTCCGATTTAGTTGGTGAATTTATGAGATTTGATGATGATGATGAT 960  
Db |||||  
Qy 764 ACCAAGAGGTTAGGTTGATGATCTTCTGGCGGCTGATGATGATGATGATGATGATGAT 808  
Db |||||

Db 961 ACAAAGAGGTTCAAGTGGAGTACTTCTCAGAGAGTGGACTGA 1005  
|||  
RESULT 13  
ABX93056  
ID ABX93056 standard; cDNA; 1223 BP.  
XX  
AC ABX93056;  
XX  
DT 22-MAY-2003 (first entry)  
XX  
DE Corn chorismate mutase cDNA.  
XX  
KW Corn; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;  
KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
XX  
OS Zea mays.  
XX  
FH Key Location/Qualifiers  
FT CDS 3..998  
FT /\*tag= a  
FT /product= "Chorismate mutase"  
FT /partial  
FT /note= "No start codon shown"  
XX  
PN US2002184658-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 03-DEC-1999; 99US-00454279.  
XX  
PR 04-DEC-1998; 98US-0110845P.  
XX  
PA (FALC/) FALCO S C.  
PA (FAMO/) FAMODU O O.  
PA (LEBJ/) LEE J.  
PI Falco SC, Famodu OO, Lee J;  
XX  
XX WPI; 2003-328651/31.  
XX P-PSDB; ABU08092.  
XX  
XX New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
XX nucleic acids, useful for facilitating design and/or identifying  
XX inhibitors of those enzymes that may be used as herbicides and for  
XX producing antibodies.  
XX  
XX Claim 3; Page 19; 32pp; English.  
XX  
XX The invention discloses isolated polynucleotides encoding chorismate  
XX mutase polypeptides. Aromatic amino acids must be included in the diets  
XX of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
XX converted to anthranilate during tryptophan biosynthesis and is converted  
XX to prephenate, the branch point for tyrosine and phenylalanine  
XX biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
XX prephenate. Also disclosed are methods for selecting an isolated  
XX polynucleotide that affects the level of expression of a tyrosine  
XX biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
XX fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
XX evaluating compounds for their ability to inhibit the activity of a  
XX tyrosine biosynthetic enzyme. The polypeptides can be used produce  
XX antibodies. Chorismate mutase and prephenate dehydrogenase are good  
XX targets for herbicides that will not affect animals, and overexpression  
XX of these enzymes may be used to increase the content of aromatic amino  
XX acid in food crops. The polypeptides may also be used to design and/or  
XX identify inhibitors of those enzymes that may be used as herbicides. The  
XX nucleic acids may be used to create transgenic plants, as probes for the  
XX genetic and physical mapping of the genes and as markers for traits  
XX linked to those genes. The sequence presented is a corn chorismate mutase  
XX cDNA





QY 332 AACAGCTTTTTCATCCTGGTGGTCTCAATTAAACATAAACAGTCCATCTGGAATG 391  
 Db 130 CCAAGGTTTTGACCCCTTTGCTTCATTTGGTCTGTGAATGATCAATATGGAAGATG 189  
 QY 392 TATTTCAAGAGATTACTTCCATTCGTTGTACTTTCGGGTGATCGGCAACTATCGCAA 451  
 Db 190 TATTTCAATGAATGTACCAATATTCACTCGGATGGCGATGATGGCAACTATGCAGAA 249  
 QY 452 ACTGAGCTAAATGACCTTTCAATTTTGCAGTCCATCTTAGAAGATTACATATGGAAG 511  
 Db 250 ACAGTTGCAATTAGATTGTCATGTCGAGGCTCTCAAGAAGAAATTCATTTGGGCAA 309  
 QY 512 TTTGTAGCTGAGTGAATTCAGGATGCTCCTCAAGACTACGAGCTTTAATTCGAGCT 571  
 Db 310 TATGTTGCTGAGTGAATTCAGGATGCTCCTCAAGACTACGAGCTTTAATTCGAGCT 369  
 QY 572 AAGGATAAAGAAGGATGATGAATTTGTCAGTCAATTTAAGCGTTGAAGACGGTGAGG 631  
 Db 370 AAGGACACTAATGCTCTGATGAATTTACTAAGCTTCAAGCTTCAAGGCTTTGAAGAAAGGCTCAAG 429  
 QY 632 AAGAGAGTTGAAGAAGGCTGTGTGTTGGCAGGAAAGTGAATCT--TAACAGTGAT 688  
 Db 430 AAGAGAGTAGAAGAAGGCAAGCAGTATTTGGACAGAAATGCTACTCTGAGGACAGTGTA 489  
 QY 689 GACAATGACAATG-----AAAACCGTAAATTTGATCCATCATGCTGCTCT 733  
 Db 490 GGCAGCAAGATGGTGATGCTGTGACAGTCACTGTAAAGTTGATCCAAAGAGTCTTCT 549  
 QY 734 AGCTTTGACAAAATTTGGGTGATACCTCTCACCAGGAGGTTTCAGGTTGAGTACCTCTTG 793  
 Db 550 AAGCTATATGATGTGGGTAAATGTCCTTTAAGCAAGGATGTTGAAGTCGAATATCTTCTC 609  
 QY 794 CGCGCTCTAGACTGA 808  
 Db 610 CGCGCTCTGACTGA 624  
 RESULT 15  
 ABA91395  
 ID ABA91395 standard; cDNA; 1217 BP.  
 AC ABA91395;  
 XX  
 DT 08-APR-2002 (first entry)  
 XX  
 DE Arabidopsis chorismate mutase-3 cDNA.  
 XX  
 KW Chorismate mutase-3; herbicide; plant; enzyme; gene; ss.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200202798-A2.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 22-JUN-2001; 2001WO-US020104.  
 XX  
 PR 05-JUL-2000; 2000US-00610040.  
 XX  
 PA (PARA-) PARADIGM GENETICS INC.  
 XX  
 PI Boyes DC, Davis KR, Woessner JP, Goriach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 PI Kjemtrup S;  
 XX  
 DR WPI; 2002-154754/20.  
 XX  
 XX Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.  
 XX  
 PS Claim 37; Page 65; 73pp; English.  
 XX

The present sequence is that of a full-length cDNA encoding Arabidopsis  
 thaliana chorismate mutase-3 (CM), an enzyme that catalyses the  
 conversion of chorismate to prephenate in a branch of the shikimate  
 pathway leading to phenylalanine and tyrosine biosynthesis. CM is  
 essential for plant growth; inhibition of CM gene expression in plant  
 seedlings results in severe chlorosis, reduced growth and developmental  
 abnormalities. The invention provides compositions for the modulation of  
 plant growth or development using CM antisense (see ABA91401) and sense  
 polynucleotides, double-stranded RNA and ribozymes, and related  
 expression cassettes and vectors. The compositions are particularly  
 useful for the modulation and inhibition of plant growth, and for  
 generating male sterile plants or plants that produce seedless fruit. CM  
 is also a target for the identification of herbicides, and methods are  
 provided for using CM in the identification of herbicide candidates.  
 Methods are also provided for identifying compounds that stimulate the  
 expression or function of CM. These can be used to promote growth and  
 development.

Query Match 20.7%; Score 211.4; DB 6; Length 1217;  
 Best Local Similarity 56.5%; Pred. No. 3.3e-47;  
 Matches 414; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
 SQ Sequence 1217 BP; 349 A; 273 C; 235 G; 360 T; 0 U; 0 Other;

CC 81 TGAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTCATTTAGAGAGCCCA 140  
 CC 363 TTAGACACTCTTTGATTCGTCAAGAGGACAGTATTATCTTTAATCTTCTTTGAACGAGCTC 422  
 QY 141 AGTTCCCTAGCAATTTCTCACACCTATGATGAAAAGTATGCTCAAAATCCAGGGTTTTTGTG 200  
 Db 423 AGTATCCTCTACACGCTGATCTATTGACGAGATGCCITTTACTATGGAAGGGTTTCAAG 482  
 QY 201 GCTCAATTTGGTGAATTTGTTTGAAGAAATACAGAGGCCATTTCAAGTAAGCTTGGAGAT 260  
 Db 483 GATCTTTAGTTGAGTTTATGCTCAGAGAAACTGAAAAGCTTTCACGAAAGGTGGACAGGT 542  
 QY 261 ACAAAACCTGAGAAACCGCTTCTTCCAGAAAATTTTACCACCATCAATTTGTGCCAT 320  
 Db 543 ACAAGATCCTGATGAGCATCCCTTTTCCCAAGATGCTTGTGCTGAACCTATCTCTTC 602  
 QY 321 CTTACTCCTTCAACACAGTTTTTGCATCTCTGCTGCTTCAATTAACAATAACAAGTCCA 380  
 Db 603 CTATTCATATCCACACAGTTTTTGCATCTGTCGCGCAATCGATAAAACATCAACAAGAG 662  
 QY 381 TCTGAAAATGATTTTCAAGAGATTACTTCCATTTGCTGCTACTTCGGGTGATGATGGCA 440  
 Db 663 TGTGGAATATGATTTTCAAAACACCTTCTCCCAAGCTGGTCAAGCCAGGGATGACGGTA 722  
 QY 441 ACTATGCGCAAACTGACAGCTAATGACCTTTTCATTTATGCACTCCATCTCTAGAAAGATTC 500  
 Db 723 ATTGTGTTTCAAGCTGCTCTCTGTGACACAATGTTGTTGACAGATCTTCAAGAGAAATTC 782  
 QY 501 ACTATGAAAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGCTCAGAGCTT 560  
 Db 783 ACTTGCCTAAATTTGTTGCTGACGCCAAGTTTCGTGAAAATCTCTGCTGCTATGAACAG 842  
 QY 561 TAATTTCAGCTAAGGATAAAGAAGGATTGATGAAATTTGTTGACATTTTACAAGGCTTGAAG 620  
 Db 843 CTATCAAGAAACAGACCCGACACAGCTGATGCAACTTCTTAACGTACGAAACGGTTGAAG 902  
 QY 621 AGACGGTGAGGAAGAGAGTTGAAAAGAGCTGTGGTGTGTTGGCAGGAAGTGAATCTTA 680  
 Db 903 AAGTAGTCAAGAGAGAGATTGAGATCAAGCCAGAAATTTTGTGCAAGCATTAACGATTA 962  
 QY 681 ACAGTGATGACAATGACAATGA---AAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCT 737  
 Db 963 ACGACCCAGAACTGAGCTGATCTCTTCAAAATACAACTAGCTTAGTTGCAAAAC 1022  
 QY 738 TGTACAAAATTTGGGTGATACCTCTCAACAGAGGTTTCAAGTTGAGTACTCTTGGGCC 797  
 Db 1023 TCTATGGAGAAAGGATCATGCCCTCTCAAAAGGAAGTCCAAATTTGAGTACTTCTGTAGAA 1082  
 QY 798 GTCTAGACTGAAG 810



Db 1083 GACTGGATTAAATG 1095

Search completed: September 24, 2004, 23:23:37  
Job time : 482 secs



Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	1020	100.0	1020	4	US-09-454-279-15	Sequence 15, Appl	
2	465.6	45.6	535	4	US-09-454-279-5	Sequence 5, Appl	
3	337.2	32.1	1006	4	US-09-610-040-3	Sequence 3, Appl	
C	337.2	32.1	1006	4	US-09-610-040-9	Sequence 9, Appl	
5	257.4	25.2	1231	4	US-09-454-279-13	Sequence 13, Appl	
6	254.8	25.0	1207	4	US-09-610-040-2	Sequence 2, Appl	
C	254.8	25.0	1207	4	US-09-610-040-8	Sequence 8, Appl	
7	234.8	23.5	1207	4	US-09-610-040-8	Sequence 8, Appl	
8	239.4	23.5	1223	4	US-09-454-279-11	Sequence 11, Appl	
9	229.4	22.5	780	4	US-09-454-279-17	Sequence 17, Appl	
10	211.4	20.7	1217	4	US-09-610-040-4	Sequence 4, Appl	
C	211.4	20.7	1217	4	US-09-610-040-10	Sequence 10, Appl	
12	206.2	20.2	579	4	US-09-454-279-7	Sequence 7, Appl	
13	98.8	9.7	5176	4	US-09-610-040-6	Sequence 6, Appl	
14	73.6	7.2	5099	4	US-09-610-040-5	Sequence 5, Appl	
15	69.2	6.8	552	4	US-09-454-279-1	Sequence 1, Appl	
16	58.4	5.7	601	4	US-09-454-279-3	Sequence 3, Appl	
C	51.8	5.1	7218	1	US-08-232-463-14	Sequence 14, Appl	
17	51.8	5.1	7218	1	US-08-232-463-14	Sequence 14, Appl	
18	38	3.7	3001	4	US-09-539-333D-151	Sequence 151, Appl	
19	38	3.7	3001	4	US-09-539-333D-155	Sequence 155, Appl	
20	37.8	3.7	2013	4	US-09-536-196-3	Sequence 3, Appl	
21	37.6	3.7	8093	4	US-10-204-708-32	Sequence 32, Appl	
22	36.2	3.5	7664	4	US-10-204-708-84	Sequence 84, Appl	
C	36.2	3.5	580073	4	US-08-545-5280-1	Sequence 1, Appl	
23	36.2	3.5	580073	4	US-08-545-5280-1	Sequence 1, Appl	
24	36	3.5	1894	3	US-09-004-731-29	Sequence 29, Appl	
C	36	3.5	1894	3	US-09-004-731-31	Sequence 31, Appl	
25	36	3.5	1894	3	US-09-032-215-3	Sequence 3, Appl	
26	36	3.5	1894	3	US-08-749-695-29	Sequence 29, Appl	
27	36	3.5	1894	3	US-08-749-695-29	Sequence 29, Appl	

361 AATTAAACATAAACAAGTCCATCTGGAAATGTATTTCAAAGAGTTACTTCCATTTGCTTGC 420  
 361 AATTAAACATAAACAAGTCCATCTGGAAATGTATTTCAAAGAGTTACTTCCATTTGCTTGC 420  
 421 TACTTCGGGTGATGATGGCAACTATGCGCAAACTGAGCTTAATGAGCTTTCATTATTGCA 480  
 421 TACTTCGGGTGATGATGGCAACTATGCGCAAACTGAGCTTAATGAGCTTTCATTATTGCA 480  
 481 GTCCATCTCTAGAAGGATTCACATATGGAAGTGTGAGCTGAGCTGAAATTCAGGATGC 540  
 481 GTCCATCTCTAGAAGGATTCACATATGGAAGTGTGAGCTGAGCTGAAATTCAGGATGC 540  
 541 TCCTCAAGACTAGAGCTTTAATTCGAGCTTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 600  
 541 TCCTCAAGACTAGAGCTTTAATTCGAGCTTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 600  
 601 GACATTTACAAGGCTTGAAGAGACGGTGAAGAGAGAGTGAAGAGAGAGTGAAGAGAGAGT 660  
 601 GACATTTACAAGGCTTGAAGAGACGGTGAAGAGAGAGTGAAGAGAGAGTGAAGAGAGAGT 660  
 661 TGGCAGAGAGTGAATCTTAACAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 720  
 661 TGGCAGAGAGTGAATCTTAACAGTGAATGAATGAATGAATGAATGAATGAATGAATGAAT 720  
 721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGT 780  
 721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGT 780  
 781 TGAGTACCTCTGGCGGCTAGACTGAAGGCAATTAAGTGAATTAAGTGAATTAAGTGAAT 840  
 781 TGAGTACCTCTGGCGGCTAGACTGAAGGCAATTAAGTGAATTAAGTGAATTAAGTGAAT 840  
 841 TGAAGATGAATATGATGTTGTTGTTCAATGATTAAGTGAATTAAGTGAATTAAGTGAAT 900  
 841 TGAAGATGAATATGATGTTGTTGTTCAATGATTAAGTGAATTAAGTGAATTAAGTGAAT 900  
 901 CAACCTTCATGCTGTTGATTCAGAAATATTTATGATGATGATGATGATGATGATGATGAT 960  
 901 CAACCTTCATGCTGTTGATTCAGAAATATTTATGATGATGATGATGATGATGATGATGAT 960  
 961 TTTTCATGCTTTTATGAGAAAGGCAATTTATATATCATGATCTTCAAAAAAATTTTAT 1020  
 961 TTTTCATGCTTTTATGAGAAAGGCAATTTATATATCATGATCTTCAAAAAAATTTTAT 1020

RESULT 2  
 US-09-454-279-5  
 ; Sequence 5, Application US/09454279  
 ; Patent No. 6627798  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Famodu, Omolayo O.  
 ; APPLICANT: Lee, Jian-Ming  
 ; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 ; FILE REFERENCE: B11299 US NA  
 ; CURRENT APPLICATION NUMBER: US/09/454,279  
 ; CURRENT FILING DATE: 1999-12-03  
 ; EARLIER APPLICATION NUMBER: 60/110,845  
 ; EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 5  
 ; LENGTH: 525  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (432)  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (514)  
 ; FEATURE:  
 ; NAME/KEY: unsure

Query Match 45.6%; Score 465.6; DB 4; Length 525;  
 Best Local Similarity 95.4%; Pred. No. 1.9e-114;  
 Matches 499; Conservative 0; Mismatches 22; Indels 2; Gaps 2;  
 8 AGCTTAATTTCAAGATGGCCAAAGCAGCAGAAACAAAGTCTGATTTCTGGCAATGTGTAC 67  
 1 AGCTTAATTTCAAGATGGCCAAAGCAGCAGAAACAAAGTCTGATTTCTGGCAATGTGTAC 60  
 68 AGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGTTCTC 127  
 61 AGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGTTCTC 120  
 128 ATTGAGAGAGCAGGATTTCCCTAGCAATTTCTCACACCTTATGATGAAAAGTATGCTCAAATC 187  
 121 ATTGAGAGAGCAGGATTTCCCTAGCAATTTCTCACACCTTATGATGAAAAGTATGCTCAAATC 180  
 188 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTTAAAGAAATACAGAGGCCATTCAGCT 247  
 181 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTTAAAGAAATACAGAGGCCATTCAGCT 240  
 248 AAGCTCGAAGATACAAAACCCCTGAAAGAAACGCTTCTCCAGAAAATTTTACCACCA 307  
 241 AAGCTCGAAGATACAAAACCCCTGAAAGAAACGCTTCTCCAGAAAATTTTACCACCA 300  
 308 TCAATTTGCGATCTTACTCTTCAACACAGTTTTTGTGATCTGCTGCTTCAATTAAC 367  
 301 TCAATTTGCGATCTTACTCTTCAACACAGTTTTTGTGATCTGCTGCTTCAATTAAC 360  
 368 ATAACAGTCCATCTGGAATATTTCAAGAGTACTTCCATGCTGCTGCTGCTGCTGCTGCT 427  
 361 ATAACAGTCCATCTGGAATATTTCAAGAGTACTTCCATGCTGCTGCTGCTGCTGCTGCT 420  
 428 GGTGATGATGGCAACTATGCGCAAACTGCGAGCTTAATGACCTTTCAATTTGCAAGTCCATC 487  
 421 GGTGATGATGGCAACTATGCGCAAACTGCGAGCTTAATGACCTTTCAATTTGCAAGTCCATC 479  
 488 TCTAGAAGGATTCACATGGAAGATTTGAGCTGAGGTGAAT 530  
 480 TCTAGAAGGATTCACATGGAAGATTTGAGCTGAGGTGAAT 521

RESULT 3  
 US-09-610-040-3  
 ; Sequence 3, Application US/09610040  
 ; Patent No. 6465217  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Boyes, Douglas  
 ; APPLICANT: Davis, Keith  
 ; APPLICANT: Woessner, Jeffrey  
 ; APPLICANT: Gorlach, John  
 ; APPLICANT: Hamilton, Carol  
 ; APPLICANT: Hoffman, Neil  
 ; APPLICANT: Klotz, Andreas  
 ; APPLICANT: Zayed, Adel  
 ; APPLICANT: Ascenzi, Robert  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
 ; FILE REFERENCE: 9128.14  
 ; CURRENT APPLICATION NUMBER: US/09/610,040  
 ; CURRENT FILING DATE: 2000-07-05  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 1006  
 ; TYPE: DNA  
 ; ORGANISM: Arabidopsis thaliana  
 ; US-09-610-040-3  
 Query Match 32.1%; Score 327.2; DB 4; Length 1006;  
 Best Local Similarity 64.2%; Pred. No. 1.5e-77;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;  
QY 22 AATGGCCAAAGCAGCAGCAAAAGTCTGATCTCGGAATGTGTACACGCTAGCTTCTGT 81  
DB 69 AAGAGTCTTCCGAATCGGATTCGGTTCGATGTCTTCAATGTACTGAGTCTTCACTTAAT 128  
QY 82 GAGAGGATTTGGTATAGGCAAGAGGATACCATTTATGTGCTCATTTGAGAGGCCAA 141  
DB 129 CAGAGATCGTTGATAGGCAAGAGACACATCGTTCAGCTTGTATCGAGAGACTAA 188  
QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAGAAAGTATGTCTCAATCCAGGCTTTTGTGG 201  
DB 189 GTTCCCACTCAATCTCTGCTTCGAGGAATCTCGTGTCTAGATTTCTGGAATTTCTC 248  
QY 202 CTGATTTGGTGAATTTCTGTTAAGATACAGAGGCCATTCAGCTAAGCTCGAAGATA 261  
DB 249 TTCTCTCACTGATTTTCTGACAGACAGAAATCATCCAGCTTAAGTAGAAGATA 308  
QY 262 CAAAAACCCCTGAAGAAACGCTTTCTCCAGAAAAATTTACCAACCATCAATTTGCCATC 321  
DB 309 TGAATACCCGGAAGAGAAATCTTTCTTCTTGAACATTCCTCACTCGGTTTTTCTAC 368  
QY 322 TTACTCCTTCAACAGTTTTTGCATCTGGTGTCTTCAATTAACATTAACAGTCCAT 381  
DB 369 GCACAAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTAAACATTAACAAACAAAT 428  
QY 382 CTGGAATATGATTTCAAGAGATTACTTCCATTGCTTGTCTACTTGGGTGATGATGCCAA 441  
DB 429 CTGGGATATTACTTTAAGAAATTTGCTTCTCTTGTTCGAACTGCGGATGATGCCAA 488  
QY 442 CTATGCGAAACTCGCAGTAAATGACCTTTCAATTTGACAGTCCATCTCTAGAAGATTCA 501  
DB 489 CTATCCATCAACTCTGCTAGTATCTCGCTGTTTCAAGCTCTTTTCGAGAGGATTCA 548  
QY 502 CTATGGAAAGTTTCTAGCTAGGTGAAATTCAGGGATGCTCTCAAGACTACGAGCCTTT 561  
DB 549 CTAGCGTAAATTTGCTAGCTAGGTCAAATTCAGAGATGCTCCACAGATTAACGAGCTGC 608  
QY 562 AATTCGAGCTAAGGATAAGAGGATTCATGAATTTGTTGACATTTCAAGCGTTGAAGA 621  
DB 609 GATTTCGGCTCAGGATAGAGAGGCTTTGATGAGCTGTGACGTTTGAGAAAGTAGAAGA 668  
QY 622 GACGCTGAGGAGAGCTTGAAGAGAGGCTGTGCTGTTTGGCGAGGAGTGAATCTTAA 681  
DB 669 AATGGTTAAGAGAGAGTGCAGAGAGAGCAGAAACGTTTGGCAAGAGTAAATTCAA 728  
QY 682 -----CAGTGATGACATCAATGAATAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
DB 729 CTCTGGCTATGGCATGAGATGAAGAGAGTATAAAGTGGATCCATTGCTTGCCTCTCG 788  
QY 736 CTGTGTACAAAATTTGGTGTATACCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGG 795  
DB 789 CATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG 848  
QY 796 CCGTCTAGACTGAA 809  
DB 849 TCGTCTCGATTGAA 862

## RESULT 4

US-09-610-040-9/c

Sequence 9, Application US/09610040

Patent No. 6465217

GENERAL INFORMATION:

APPLICANT: Boyes, Douglas

APPLICANT: Davis, Keith

APPLICANT: Woessner, Jeffrey

APPLICANT: Goriach, John

APPLICANT: Hamilton, Carol

APPLICANT: Hoffman, Neil

APPLICANT: Kloti, Andreas

APPLICANT: Zayed, Adel

APPLICANT: Ascenzi, Robert

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHESIS;  
TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS  
FILE REFERENCE: 9128.14  
CURRENT APPLICATION NUMBER: US/09/610,040  
CURRENT FILING DATE: 2000-07-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 9  
LENGTH: 1006  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-610-040-9

Query Match 32.1%; Score 327.2; DB 4; Length 1006;

Best Local Similarity 64.2%; Pred. No. 1.5e-77;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGGCCAAAGCAGCAGCAAAAGTCTGATCTCGGAATGTGTACACGCTAGCTTCTGT 81  
DB 938 AAGAGTCTTCCGAATCGGATTCGGTTCGATGTCTTCAATGTACTGAGTCTTGAATTAAT 879  
QY 82 GAGAGGATTTGGTATAGGCAAGAGGATACCATTTATGTGCTCTCATTTGAGAGGCCAA 141  
DB 878 CAGAGATCGTTGATAGGCAAGAGACACCATCGTCTTCAGCTTGTATCGAGAGACTAA 819  
QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAGAAAGTATGTCTCAATCCAGGCTTTTGTGG 201  
DB 818 GTTCCCACTCAATCTCTGCTTTCGAGGAATCTCGTGTCTAGATTTCTGGAAGTTTCTC 759  
QY 202 CTCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGGCTGGGAAGATA 261  
DB 758 TTCTCTCACTGATTTTTCGTCAGAGACAGAAATCATCCAGCTTAAGTAGAAGATA 699  
QY 262 CAAAAACCCCTGAAGAAACGCTTTCTCCAGAAAAATTTACCAACCATCAATTTGCCATC 321  
DB 698 TGAATACCCGGAAGAGAAATCTTTCTCTTGTGAGAACATTCCTCACTCGGTTTTTCTAC 639  
QY 322 TTACTCCTTCAACAGTTTTTGCATCTGGTGTCTTCAATTAACATTAACAGTCCAT 381  
DB 638 GCACAAAATATCCATCGGCTTTTCACCCCTAAGCTCTATCTGTTAATTAACAAACAAAT 579  
QY 382 CTGGAATATGATTTTCAAGAGATTACTTCCATTGCTTGTCTACTTGGGTGATGATGGCAA 441  
DB 578 CTGGGATATTTACTTTAAGAAATTTGCTTCTTGTTCGAAACCTGCGCATGATGGCAA 519  
QY 442 CTATGGCAAACTCGCAGCTAATGACCTTTCAATTTGAGTCCATCTCTAGAAGGATTCA 501  
DB 518 CTATCCATCAACTGCTGTAGTGTCTCGCTGTTTACAGCTCTTTTCGAGAGGATTCA 459  
QY 502 CTATGGAAGTTTGTAGCTAGGTGAAATTCAGGGATGCTCTCAAGACTACGAGCCTTT 561  
DB 458 CTACGTAATTTTGTAGCTGAGTCAAATTCAGAGATGCTCCACAGATTAACGAGCTGC 399  
QY 562 AATTCGAGCTAAGGATAAGAGGATTTGATGAATTTGTTGACATTTACAGGCTTTGAGA 621  
DB 398 GATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAAGAGTGAAGA 339  
QY 622 GACGCTGAGAGAGAGATTGAAAGAGGCTGTGCTGTTTGGCGAGGAGTGAATCTTAA 681  
DB 338 AATGGTTAAGAGAGAGTGCAGAGAGAGCAAGAAAGTTTGGCAAGAGTAAATTCAA 279  
QY 682 -----CAGTGATGACATCAATGAATAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735  
DB 278 CTCTGGCTATGGCATGAGAGTAAAGAGAGTATAAAGTGGATCCATTGCTTGCCTCTCG 219  
QY 736 CTGTGTACAAAATTTGGTGTATACCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGG 795  
DB 218 CATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTTCTACG 159  
QY 796 CCGTCTAGACTGAA 809  
DB 158 TCGTCTCGATTGAA 145

## RESULT 5

US-09-454-279-13

; Sequence 13, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; FILE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; TITLE OF INVENTION: US/09/454,279

; CURRENT APPLICATION NUMBER: 1299 US NA

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 13

; LENGTH: 1231

; TYPE: DNA

; ORGANISM: Oryza sativa

US-09-454-279-13

## Query Match

Best Local Similarity 25.4%; Score 257.4; DB 4; Length 1231;

Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;

Qy 24 TGCCCAAGCAGCAGCAAGCTCTGTTCTGGGAATGTTACAGCGTAGCTTCTGTCA 83  
Db 239 TGACCAAGGAGGAGCAACAGAGATAGATCAAGTGAATATGACCTTGGCAACATTA 298  
Qy 84 GAGAGGATTTGTTAGGCAAGAGATACCATCATTTATGTTCTATTTGAGAGAGCAAGT 143  
Db 299 GAACCTCTTGGTTAGGCAAGAGACAGCATCATATTCAGCGCTCTTAGAGAGACACAGT 358  
Qy 144 TCCTAGCAATTCACACCTATGATGAAAGATGCTCAAAATCCAGGTTTTTGGCT 203  
Db 359 TTGCTACAAATGCTGATATATATGATGAAAGATGCTTTCATGTTGATGATTTGATGCT 418  
Qy 204 CATTTGTTGAATTTGTTTGAATATACAGAGGCAATTCAGCTAAGGCTGGAAGATACA 263  
Db 419 CTTTGTGTAATTCATGTTAGAGAAACCGAAACCTACATCAACAGGTTGGAGATACA 478  
Qy 264 AAAACCTGAGAAAGAGGCTTTCCAGAGAAATTTACCATCAATTTGGCATCTT 323  
Db 479 AGAGCCCTGATGAGCAACCAATTTCTCCGAGGATCTGCTGAACCACTGTTGCCACCTC 538  
Qy 324 ACTCTTTCAACAGTTTGTGATCTCTGCTGCTGCTTCAATTAACATAAAGAGTCCATCT 383  
Db 539 TCAGTATCCAGGTTTGTGATCTTATGCTGATCTTATTAATATCAACAGGAGATT 598  
Qy 384 GGAATATGATTTCAAGAGTACTTCCATGCTTCTGCTACTCTGCGGTGATGATGCAACT 443  
Db 599 GGAATATGATTTTGTGATGAGCTTCTTCCAGATTTAGTGAAGAGGAAAGTGAAT 658  
Qy 444 ATGCGCAAACTGAGCTAATGACCTTTCATTTATGAGTCCATCTCTAGAGGATTCAT 503  
Db 659 ATGATTCAGTCTCTTTGTGACAGATCTGTTGAGCGGCTCTCCAAAGAAATTCAT 718  
Qy 504 ATGGAAGTTTGTAGTGAAGTGAATTCAGGATGCTCTCAAGACTACGAGCTTTAA 563  
Db 719 ATGGAAGTTTGTGAGAGGCTAAGTTTCAAGAGTCTCTCAAGCTTATACATGCTGGA 778  
Qy 564 TTCGAGCTAAGATTAAGAGGATTCAGTAATTTGATGATTTTACAGGTTGAGAGA 623  
Db 779 TAATGACAGAGCTGCGATCACTTAATGACCTCTCTACCTATGAAACGGTGGAGCGT 838  
Qy 624 CGGTGAGGAGAGTTGAAAGAGGCTGTTGGTGTGTTGGGAGGAGTGAATCTTAACA 683  
Db 839 CTATTGAACATAGGTTGGAAGCTAAGGCTAAGATCTTTGGAAGAGGATTTAGCG 898  
Qy 684 GTGATGACATGACAT---GAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGT 740

Db

899 CTGAAGACACGGCGCTCCACCAATGTACAAGATAGGCCAGTTTGGTGGCTGAAGTGT 958

Qy

741 AAAAAATTTGGTGATACCTCTCAACAGAGGTTTCAAGTTGATGATCTCTTGGCGGTC 800

Db

959 ACAGCTACAGGATCATGCGCTAACCAAGAGGTTCAAGTAGCTACTTCTGAGGAGAT 1018

Qy

801 TAGACTGAAGCATTACAATGCA 823

Db

1019 TGGATTGATTGTTTACGATTGTA 1041

## RESULT 6

US-09-610-040-2

; Sequence 2, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woessner, Jeffrey

; APPLICANT: Gornach, John

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE

; FILE REFERENCE: 9128.14

; CURRENT APPLICATION NUMBER: US/09/610,040

; CURRENT FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1207

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-610-040-2

## Query Match

Best Local Similarity 25.0%; Score 254.8; DB 4; Length 1207;

Matches 494; Conservative 0; Mismatches 342; Indels 24; Gaps 1;

Qy 68 ACCTAGCTTCTGTGAGAGGATTTGGTTAGCAGAGGATACCATCTTTATGCTCTC 127  
Db 253 ACTCTTGAAGGATTTAGAACTCTTTGATCCGCTCAAGAGGACAGCAATTATTTGGGTA 312  
Qy 128 ATTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAAAATC 187  
Db 313 TTGAGAGAGCCCAAGTACTGTTACATGCTGATACCTTATGATCTCTGCTTTGACATG 372  
Qy 188 CAGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTGTAAGAAATACAGAGGCCATTCAAGCT 247  
Db 373 GATGGTTTCAATGGTTCTTTGGTTGATGATGTTTAAAGGCACTGAGAAGCTTCAAGCT 432  
Qy 248 AAGCTGGAAGATACAAAACCCCTGAAGAAAACCGCTTCTCCAGAAAATTTACCACCA 307  
Db 433 AAGTTTGTAGGTTTAAAGTCTCTGATGACATCTTTTCTCCCTGATGATCTACAGAG 492  
Qy 308 TCAATTTGTCATCTTACTCTCTTCAACAGATTTTTCATCTCTGCTGCTGCTCAATTAAC 367  
Db 493 CCTATGTTGCTCTCTCTTCAAGTACCCAAAGGTTTGTCAATTTGCTGCTGATTCGATAAC 552  
Qy 368 ATAAAAGGCTCATCTGGAATGTTTCAAGAGTACTTCCATTTGCTGCTCTCTGCT 427  
Db 553 ATAAAAGGATATGGAACATGTTACTTCAAGAGCTTGTTCAGAGATTTGTGAGAAA 612  
Qy 428 GGTGATGAGGCAATATGCGCAACTGAGCTTAATGACCTTTCAATTTATGAGTCCATC 487  
Db 613 GCGCATGATGTTAACTACCGCTCAACAGCTGCTGTGACGCTATCTGCTTCAAGTCTC 672  
Qy 488 TCTAGAGGATTCACATATGGAAGTTTGTAGCTGAGTGAATTCAGGATCTCTCTCAA 547  
Db 673 TCAAGAGATTCATTTACGTTAAATTTGTGAGAGCTTAATTTCAAGCTTCAACCGAA 732





Matches 463; Conservative 0; Mismatches 351; Indels 3; Gaps 1;  
QY 2 CACGAGAGCTTAATTTCAAGAAATGGCCAAAGCAGCAGACAAAGTCTCTGATTCGGAAAT 61  
Db 189 CGCGGGCCAACTACCTCCGCAAGCCCATAGCTTAAGGTAGAGGGTGTATCGAAGTGAC 248  
QY 62 GTGTACAGCTAGCTCTGTGAGAGGAGTTTGGTTAGGCAAGAGATACCATCATTTAT 121  
Db 249 ATATTGACATTTGGATAGCATCAGACAGATTTTGAATAGACTAGAACAGACATCATATT 308  
QY 122 GGTCTCATTCAGAGAGCCAGTTCCCTAGCAATTTCTCAGACCTATGATGAAAGATGCT 181  
Db 309 GGCCTTTTGGAGAGAGCAGATTTTGTACAAATGCTGATACATATATAGCAATGCTTTC 368  
QY 182 CAAATCCAGGGTTTGTGGCTCATTTGGTGGAAATTTGTTAAGAAATACAGAGGCCATT 241  
Db 369 CACATGGATGGTTTGGAGGATCTTTGGTTGAATATATAGTTAGAGAACTGAAAGCTC 428  
QY 242 CAGCTAAGGCTGGAAGATACAAACCTCGAGAAAGCCCTCTTCCAGAAATTTA 301  
Db 429 CATGCACAGTTGGGAGATCAAGAGCCAGATGAGACCCCTTTCTTTTCCAGGATCTG 488  
QY 302 CCACCATCAATTTGGCTCATTTACTCCTTCAACACAGTTTGTGATCTGGTGTCTTCA 361  
Db 489 CTTGAGCCCGGTTGCCACCTATGCAATACCCAGGGTTTTCATCCCATTCGTGATCT 548  
QY 362 ATTAAACATAAAGCTTCATCTGGAAATGTTATTTCAAGAGTTACTTCCATGCTTGTCT 421  
Db 549 ATCAATATCAACAAAGAGATTTGGAAATGTTATTTGATGAACTTCTTCCAGATTTGGT 608  
QY 422 ACTTCGGGTGATGATGGCACTATGCGCAACTGCGCAAGTAAAGCTTTTCAATTATTCGAG 481  
Db 609 AAGAGAGGAGTGAATGATGCTGATCCAGTCTCTTTGTGACACACACTGCTTTCGAG 668  
QY 482 TCATCTCTAGAGAGTTCATATGAGAAATTTGTAGCTGAGGTGAATTCAGGATGCT 541  
Db 669 GCACCTCTCCAGAGGATCCACTATGCGAAGTTTGTGCGAGAGCCCAAGTTTCAAGATCC 728  
QY 542 CTTCAAGACTAGAGCTTTTAAATTCGAGCTAAGGATTAAGAGGATGATGAAATGTTG 601  
Db 729 CTTGAAGCTTACGCCAGGATTAATGATGATGATGATGATGATGATGATGATGATGAT 788  
QY 602 ACATTTTCAAGCGTTGAAGAGACGGTGAAGAGAGAGTTGAAAGAGGCTGTGTGTGTT 661  
Db 789 ACATATGAAACGGTGGAGCGTCTATCGAAACACAGGGTGGAGCCCAAGATCTTC 848  
QY 662 GGCAGAGAGTGAATCTTAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 718  
Db 849 GGCAGAGAGTGAATCTTAACAGTATGATGATGATGATGATGATGATGATGATGATGAT 908  
QY 719 CCATCAGTGGCTTTAGCTTTGTACAAAATTTGGGTGATACCTCTCACCAGAGAGTTTCAG 778  
Db 909 CCGAGCTTCTGCGCGAGCTGTACAGTACAGATCATGCGCTAACCCAGAGGTTTCAG 968  
QY 779 GTTGAAGTCTTCTGCGCGCTGTACAGTGAAGGCAATT 815  
Db 969 GTGCGGTACTTCTTGAAGAGGCTGATTTGATGTTGTT 1005

## RESULT 9

US-09-454-279-17

; Sequence 17, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454, 279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-454-279-17

Query Match 22.5%; Score 229.4; DB 4; Length 780;  
Best Local Similarity 63.6%; Pred. No. 1.3e-51;  
Matches 391; Conservative 0; Mismatches 206; Indels 18; Gaps 2;

QY 212 GAATTTGTTTAAAGATACAGAGGCCATTCAAGCTTAAGGCTGGAAGATACAAAAACCCCT 271  
Db 10 GAGTTCTTCGTTCCGGAAGCCGAGGTCCTGCAGCCCAAGGCTGGACACTATCAAAAGCCA 69  
QY 272 GAAGAAACCCCTTCTCCAGAAATTTACCAACCATCAATTTGTCATCTTACTCTTTC 331  
Db 70 GAAGATGTTCCATTCCTTCTCAAGATCTTCCACCACTCTCTTCTTCCAAAGGTGCG 129  
QY 332 AAACAGTTTTCATCTCTGCTGCTTCAATTAACATAAACAAGTCCCATCTGGAATAATG 391  
Db 130 CCAAGGTTTTCACCCCTTTTGTCTTCAATTTGCTGCTGATGATGCAATATGGAAGATG 189  
QY 392 TATTTCAAGAGTACTTCCATTTGCTTGTCTTCTTCTGCTGATGATGCAACTATGCGAA 451  
Db 190 TATTTCAATGAATGCTTACCATTAATTCATGCGGATGGCGATGATGCAACTATGCGAA 249  
QY 452 ACTGCAGCTAATGACCTTTTCATTTATTCAGTCCATCTCTAGAGAGGATTCATATGGAAG 511  
Db 250 ACAGTTGCTATGATTTTGCATGCTGCGAGCTCTCTCAAGAGAAATTCATTTGCGCAA 309  
QY 512 TTTGTAGCTAGGTGAAATTCAGGATGCTTCTCAAGACTACGAGCTTTAATTCAGCT 571  
Db 310 TATGTTCTGAGTGAATTTCAAGAGCGGCTTCAAGATTTATAGCCCAACCAATACGTCT 369  
QY 572 AAGGATAAAGAGGATTTGATGAAATTTGTGACATTTTCAAGCGTTCAAGAGCGGTGAGG 631  
Db 370 AAGGACTAATGCTCTGATGAATTTACTAAAGTTCACGCTGTTGAAGAAAGTCAAG 429  
QY 632 AAGAGAGTTGAAAGAGGCTGTGTTTGGCGAGAAAGTGAATCT---TAAAGTGTAT 688  
Db 430 AAGAGAGTGAAGAGAGGCAAGGATATTTGGACAGATGTCACTCTGGAGGACAGTGA 489  
QY 689 GACAATGACATG-----AAAAACGTAAATTTGATCCATCAGTGGCTTCT 733  
Db 490 GGCAGCAGATGCTGATGCTGTCAGTCACTGTAAAGTTGATCCAAAGTCTTTCT 549  
QY 734 AGCTTGTACAAAATTTGGGTGATACCTCTCACCAGAGGTTTCAGGTTGAGTACCTTTG 793  
Db 550 AAGCTATATGATGTGGGTAAATGCTTTTAAAGAGGATGTTGAAGTTCGAAATATCTTCTC 609  
QY 794 CGCCTCTAGACTGA 808  
Db 610 CGCGCTCTTGACTGA 624

## RESULT 10

US-09-610-040-4

; Sequence 4, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woessner, Jeffrey

; APPLICANT: Gorlach, Jorn

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE

; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS

FILE REFERENCE: 9128.14  
CURRENT APPLICATION NUMBER: US/09/610,040  
CURRENT FILING DATE: 2000-07-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 4  
LENGTH: 1217  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-610-040-4

Query Match 20.7%; Score 211.4; DB 4; Length 1217;  
Best Local Similarity 56.5%; Pred. No. 9.1e-47;  
Matches 414; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 81 TGAGAGAGATTGGTTAGGCAAGAGATACCATTTATGGTCTCATTTGAGAGGCCA 140  
DB 363 TTAGACACTCTTGAATCGTCAGAGGACAGTATATCTTTAATCTCTTGAAGAGCTC 422  
QY 141 AGTTCCCTAGCAATTCACACCTATGATGAAAGTATGCTCAATCCAGGGTTTTGTG 200  
DB 423 AGTATCGCTACACGCTGATCTTATGACGAGGATGCTTTACTATGGAAGGGTTTCAAG 482  
QY 201 GCTCATTTGGTGAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGSCTGGAAGAT 260  
DB 483 GATCTTTAGTTGAGTTTATGGTACAGAACTGAAAGCTTCACGAAAGGTGACAGGT 542  
QY 261 ACAAAAACCCCTGAAGAAAACGCTTTCTCCAGAAAAATTTACCACCATCAATTTGTGCCAT 320  
DB 543 ACAAGAGTCTGTATGAGCATCCCTTTTCCCAACAATGCTTGCCTGAACCTATCTCTCCTC 602  
QY 321 CTACTCTCTTCAAACAGTTTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380  
DB 603 CTATTCATACCCACAGGTTTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662  
QY 381 TCTGAAAATGATTTCAAAGAGTACTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440  
DB 663 TGTGGAATGATTTCAAACACCTTCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722  
QY 441 ACTATGCGCAATGTCAGCTAATGACCTTTTCAATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 500  
DB 723 ATTGTGTTTCACTGCTCTCTGTGACACAAATGTTTGGCAGATACCTTTCAAAGAGAAATTC 782  
QY 501 ACTATGGAAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGACTACGAGCTTACGAGCTTT 560  
DB 783 ACTTGGTAAATTTGTTGCTGACGCCAAGTTTCTGGAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842  
QY 561 TAATTCGAGCTAAGGATAAAGAGGATTTGATGAAATTTGTCACATTTTCAAGCGTTGAAG 620  
DB 843 CTATCAAGAAACAAGACCGGACACAGCTGATGCAACTTCTAACTGACGAAACGGTTTGAAG 902  
QY 621 AGACGGTGAAGAGAGTTGAAAGAGGCTGTGGTGTGTTGGCAGGAGTGAATCTTA 680  
DB 903 AAGTAGTCAAGAGAGGTTGAGATCAAGCCAGAAATTTTGTGTCAGAGCATACGATTA 962  
QY 681 ACAGTGATGACAAATGACAAATGA---AAACCGTAAATTTGATCCATCAGTGGCTTCTAGCT 737  
DB 963 ACGACCCAGAACTGAAGCTGATCTCTTCAAAAATACAACTAGCTTAGTTGCAAAAC 1022  
QY 738 TGTACAAAATTTGGTGTATCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGGGCC 797  
DB 1023 TCTATGGAGAAGGATCATGCCCCCTCAAAAAGGAGTCCAAATTTGAGTACTTGTCTAGAA 1082  
QY 798 GTCTAGACTGAAG 810  
DB 1083 GACTGGATTAATG 1095

RESULT 11

US-09-610-040-10/c

Sequence 10, Application US/09610040

Patent No. 6465217

GENERAL INFORMATION:

APPLICANT: Boyes, Douglas  
APPLICANT: Davis, Keith  
APPLICANT: Woesner, Jeffrey  
APPLICANT: Goriach, Jörn  
APPLICANT: Hamilton, Carol  
APPLICANT: Hoffman, Neil  
APPLICANT: Kloti, Andreas  
APPLICANT: Zayed, Adel  
APPLICANT: Ascenzi, Robert  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
FILE REFERENCE: 9128.14  
CURRENT APPLICATION NUMBER: US/09/610,040  
CURRENT FILING DATE: 2000-07-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 10  
LENGTH: 1217  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
US-09-610-040-10

Query Match 20.7%; Score 211.4; DB 4; Length 1217;  
Best Local Similarity 56.5%; Pred. No. 9.1e-47;  
Matches 414; Conservative 0; Mismatches 316; Indels 3; Gaps 1;  
QY 81 TGAGAGAGATTGGTTAGGCAAGAGATACCATTTATGGTCTCATTTGAGAGGCCA 140  
DB 855 TTAGACACTCTTGAATCGTCAGAGGACAGTATATCTTTAATCTCTTGAAGAGCTC 796  
QY 141 AGTTCCCTAGCAATTCACACCTATGATGAAAGTATGCTCAATCCAGGGTTTTGTG 200  
DB 795 AGTATCGCTACAAACGCTGATCTTATGACGAGGATGCTTTACTATGAAAGGGTTTCAAG 736  
QY 201 GCTCATTTGGTGAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGSCTGGAAGAT 260  
DB 735 GATCTTTAGTTGAGTTTATGGTACAGAACTTGAAGAGCTTCAAGCAAGGTGACAGGT 676  
QY 261 ACAAAAACCCCTGAAGAAAACGCTTTCTCCAGAAAAATTTACCACCATCAATTTGCGCAT 320  
DB 675 ACAAGAGTCTGTATGAGCATCCCTTTTCCCAACAATGCTTGCCTGAACCTATCTCTCCTC 616  
QY 321 CTACTCTCTTCAAACAGTTTGTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 380  
DB 615 CTATTCAATACCCACAGGTTTGTGATCGTTGCGCGAATCGATAAACAACAAGAGG 556  
QY 381 TCTGGAATGATTTCAAAGAGTTACTTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 440  
DB 555 TGTGGAATGATTTCAAACACCTTCTCCCAAGCTGCTCAAGCCAGGGATGACGTA 496  
QY 441 ACTATCGCAAACTGACGCTAATGACCTTTCAATTTATGAGTCCATCTCTAGAAAGATTC 500  
DB 495 ATTGTTTCACTGCTCTCTGTGACACAAATGTTTGCAGATACTTTCAAGAGNAATTC 436  
QY 501 ACTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCCTCAAGACTACGAGCTTT 560  
DB 435 ACTTGGTAAATTTGTTGCTGACGCAAGTTTCTGTAAGAAATCTCTGCTGCTGCTGCTGCTGCT 376  
QY 561 TAATTCGAGCTAAGGATAAAGAGGATTTGATCAATTTGTCACATTTTCAACGCTTTGAAG 620  
DB 375 CTATCAAGAAACAAGACCGGACACAGCTGATGCAACTTCTAACGTAACGAAACGGTTGAAG 316  
QY 621 AGACGGTGAAGAGAGTTGAAAGAGGCTGTGGTGTGTTGGGAGGAGTGAATCTTTA 680  
DB 315 AAGTAGTCAAGAGAGGTTGAGATCAAGCCAGAAATTTTGTGTCAGAGCATACGATTA 256  
QY 681 ACAGTGATGACAAATGACAAATGA---AAACCGTAAATTTGATCCATCAGTGGCTTCTAGCT 737  
DB 255 ACGACCCAGAACTGAAGCTGATCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGGGCC 196  
QY 738 TGTACAAAATTTGGGTTGATACCTCTCAACAGAGGTTTCAAGTTGAGTACCTCTTGGGCC 797  
DB 195 TCTATGGAGAAGGATCATGCCCCCTCAAAAAGGAGTCCAAATTTGAGTACTTGTCTAGAA 136

QY 798 GTCTAGACTGAAG 810  
Db 135 GACTGGATTAATG 123

## RESULT 12

US-09-454-279-7  
; Sequence 7, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 579  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (417)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (484)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (536)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (551)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (585)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (577)  
US-09-454-279-7

Query Match 20.2%; Score 206.2; DB 4; Length 579;  
Best Local Similarity 65.1%; Pred. No. 1.7e-45;  
Matches 304; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

QY 212 GAATTTGTTTGAAGATACAGAGGCCATTCAAGCTAAGCTTGGAGATACAAAACCCCT 271  
Db 3 GAGTTCCTTCGTTCCGGAAGCGGAGGTCCTGCACGCCAAGCTGGACATATCAAAAGCCA 62  
QY 272 GAAGAAAACGCTTCTTCCAGAAATTTACACCATCAATTTGCGCATCTTACTCCTTC 331  
Db 63 GAAGATGTTCCATCTTCTCTCAAGATCTTCCACCATCTGTTCTTCTACCAAGTGGC 122  
QY 332 AAACAGTTTTTGGCATCTGTTGCTGCTCAATTAACATAAAACAGTCCATCTGGAAAATG 391  
Db 123 CCAAAGTTTTTGACCCCTTTTCTTCAATTTGCTGTGATGATGCAATATGGAAGATG 182  
QY 392 TATTTCAAGAGTACTTCTCAATGCTTGTCTTCTCGGGTGAATGGCAACTATGGCAA 451  
Db 183 TATTTCAAGAGTATGTTACCAATTTTCACTCGGATGGCGATGATGGCAACTATGCGAA 242  
QY 452 ACTGCGAGCTAATGACCTTTTCATTTATGCAATCTCTAGAGGATTCACATATGGAAG 511  
Db 243 ACAGTTGCATTAGATTTTGCATGCTGCGAGGCTCTCTCAAGAAGATTCATTTGGCAA 302  
QY 512 TTTGTAGCTGAGGTGAAATTCAGGGATGCTTCTCAAGACTAGAGCCTTTATTTGAGCT 571  
Db 303 TATGTTGCTGAGGTGAAATTCAAAGACGCGCTTCAAGATTTATAGCCCAACCAATACGTCT 362

QY 572 AAGGATTAAGAGGATTGATGAATAATTTTGACATTTTCAAGCGTTGAAGAGCGGTGAGG 631  
Db 363 AAGGACACTAATGCTCTGATGACCTTACTAACGTTTACGGCTGTTGAAGAAAAGGTCAAG 422  
QY 632 AAGAGAGTTGAAAAGAGGCTGTGCTTTTGGCGAGGAAGTGAATCT 678  
Db 423 AAGAGAGTAGAGAAGAGGCAAGGATATTTGGACAGAGATGCTCACTCT 469

## RESULT 13

US-09-610-040-6  
; Sequence 6, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 6  
; LENGTH: 5176  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-610-040-6

Query Match 9.7%; Score 98.8; DB 4; Length 5176;  
Best Local Similarity 65.1%; Pred. No. 1.3e-16;  
Matches 164; Conservative 0; Mismatches 82; Indels 6; Gaps 1;

QY 564 TTCGAGCTAAGGATTAAGAGGATTGATGAATTTTGACATTTTCAAGCGTTGAAGAGA 623  
Db 3869 TTCGACTGTAGGATAGAGGCTTTGATGAAGCTGTGACGTTTGAGAAAGTAAAGAA 3928  
QY 624 CGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTTTGGCGAGGAAGTGAATCTTAA-- 691  
Db 3929 TGGTTAAGAGAGAGAGTGCAGAGAAAGCAGTGTGGACAAGTAAATTCAACT 3988  
QY 582 ----CAGTGATGACATGACAAATGAACCGTAAATTTGATCCATCAGTGGCTCTAGCT 737  
Db 3989 CTGCTATGCGATGAGAGTAAGAAGATTAAGAATGATCCATTCGCTCTCGCA 4048  
QY 738 TGTAACAAAATTTGGGTGATACCTCTCACCAGAGGTTTCAAGTTGAGTACCTCTTGGCC 797  
Db 4049 TCTACGGGATGCTTATCCCTCTCACTAAGCTCGTTGAGTTGAGTATCTTCTACGTC 4108  
QY 798 GTCTAGACTGAA 809  
Db 4109 GTCTCGATTGAA 4120

## RESULT 14

US-09-610-040-5  
; Sequence 5, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil

APPLICANT: Kloti, Andreas  
APPLICANT: Zayed, Adel  
APPLICANT: Ascenzi, Robert  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
FILE REFERENCE: 9128.14  
CURRENT APPLICATION NUMBER: US/09/610,040  
CURRENT FILING DATE: 2000-07-05  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 5  
LENGTH: 5099  
TYPE: DNA  
ORGANISM: Arabidopsis thaliana  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (4049)..(4049)  
OTHER INFORMATION: "n" indicates any nucleotide (A, C, T, or G).  
US-09-610-040-5

Query Match 7.2%; Score 73.6; DB 4; Length 5099;  
Best Local Similarity 62.5%; Pred. No. 6.2e-10;  
Matches 115; Conservative 0; Mismatches 69; Indels 0; Gaps 0;  
QY 68 ACGCTAGCTTCTGTGAGAGAGATTGGTTAGGCAAGAGATACCATCATTTATGGTCTC 127  
Db 3428 ACTCTGAGGTTATAGAACTCTTTGATCCGTCAAGAGGACACATTATTTGGGCTA 3487  
QY 128 ATTGAGAGAGCAAGTTCCTTAGCAATCTCACACCTATGATGAAAAGTATGCTCAAATC 187  
Db 3488 TTGGAGAGAGCAAGTACTGTACATGCTGATCTATGATCTCTGCTTTTGACATG 3547  
QY 188 CAGGTTTTCGTGCTCATTTGGTGAATTTGTTTGAATATAGAGAGGCGCATTCAGCT 247  
Db 3548 GATGGTTTCAATGGTCTTTGGTTGAGTACATGTTAAAGCACTGAGAAGCTTCAGCT 3607  
QY 248 AAGG 251  
Db 3608 AAGG 3611

RESULT 15  
US-09-454-279-1  
Sequence 1, Application US/09454279  
Patent No. 6627798  
GENERAL INFORMATION:  
APPLICANT: Falco, S. Carl  
APPLICANT: Famodu, Omolayo O.  
APPLICANT: Lee, Jian-Ming  
TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
FILE REFERENCE: BB1299 US NA  
CURRENT APPLICATION NUMBER: US/09/454,279  
CURRENT FILING DATE: 1999-12-03  
EARLIER APPLICATION NUMBER: 60/110,845  
EARLIER FILING DATE: 1998-12-04  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: Microsoft Office 97  
SEQ ID NO 1  
LENGTH: 552  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (442)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (456)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (460)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (472)

FEATURE:  
NAME/KEY: unsure  
LOCATION: (486)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (494)..(495)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (513)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (518)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (522)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (534)  
US-09-454-279-1  
Query Match 6.8%; Score 69.2; DB 4; Length 552;  
Best Local Similarity 55.4%; Pred. No. 4.2e-09;  
Matches 134; Conservative 0; Mismatches 108; Indels 0; Gaps 0;  
QY 2 CACGAGAGCTTAATTTCAAGATGGCCAAAGCAGCAGACAAAGTCCTGATTCTGGGAT 61  
Db 189 CGCGCGGCCAACAACTCCGCAACGCCCATAGCTAAGGTAGAGAGGGTTGATCGAAGTGAC 248  
QY 62 GTGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTAT 121  
Db 249 ATATTGACATTGGATAGCATCAGACAAGTTTGTATTAGACTAGAGAGCATCATATTT 308  
QY 122 GGTCTCATTTGAGAGAGCAAGTTCCTTAGCAATTTCTCACACCTATGATGAAAAGTATGCT 181  
Db 309 GGCCCTTTTGGAGAGAGCACAGTTTGTTCACAAATGCTGATACATATGATAGCAATGCTTC 368  
QY 182 CAAATCCAGGGTTTTTGTGGCTCATTTGGTGAATTTGTTTGAAGATACAGAGGCCATT 241  
Db 369 CACATGGATGGTTTTGGAGGATCTTGGTTGAATATAGTTAGAGAACTGAGAAGCTCCATG 428  
QY 242 CA 243  
Db 429 CA 430

Search completed: September 25, 2004, 01:26:32  
Job time : 102 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 00:33:15 ; Search time 554 Seconds  
(without alignments)  
9325.358 Million cell updates/sec

Title: US-10-624-061-15

Perfect score: 1020  
Sequence: 1 GCACGAGAGCTTAATTCAGAAATGCCCAAGCAGCAGCAAAAGTCTGATTCGGAA 60

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	100.0	1020	9	US-09-454-279-15
2	1020	100.0	1020	16	US-10-624-061-15
3	985	96.6	1079	13	US-10-425-114-7639
4	940.4	92.2	1653	13	US-10-424-599-93036
5	689.4	67.6	992	13	US-10-425-114-15059
6	465.6	45.6	525	9	US-09-454-279-5
7	465.6	45.6	525	16	US-10-624-061-5
8	327.2	32.1	1006	15	US-10-267-763-3
9	327.2	32.1	1006	15	US-10-267-763-9
10	326.2	32.0	798	9	US-09-938-842A-972
11	326.2	32.0	798	11	US-09-938-842A-972
12	260.4	25.5	694	13	US-10-424-599-141578
13	257.4	25.2	1016	17	US-10-437-963-93956
14	257.4	25.2	1231	9	US-09-454-279-13

15	257.4	25.2	1231	16	US-10-624-061-13	Sequence 13, Appl
16	254.8	25.0	1207	15	US-10-267-763-2	Sequence 2, Appl
17	254.8	25.0	1207	15	US-10-267-763-8	Sequence 8, Appl
18	254.8	25.0	1423	13	US-10-424-599-123820	Sequence 123820,
19	251	24.6	1005	9	US-09-938-842A-737	Sequence 737, App
20	251	24.6	1005	11	US-09-938-842A-737	Sequence 737, App
21	239.4	23.5	1146	13	US-10-425-114-7285	Sequence 7285, Ap
22	239.4	23.5	1223	9	US-09-454-279-11	Sequence 11, Appl
23	239.4	23.3	1223	16	US-10-624-061-11	Sequence 11, Appl
24	237.8	23.3	1197	13	US-10-425-114-27576	Sequence 27576, A
25	237.2	23.3	1500	17	US-10-767-701-14353	Sequence 14353, A
26	232.4	22.8	768	17	US-10-437-963-49372	Sequence 49372, A
27	229.4	22.5	780	9	US-09-454-279-17	Sequence 17, Appl
28	229.4	22.5	780	16	US-10-624-061-17	Sequence 17, Appl
29	227.4	22.3	427	13	US-10-424-599-81826	Sequence 81826, A
30	211.4	20.7	1217	15	US-10-267-763-4	Sequence 4, Appl
31	211.4	20.7	1217	15	US-10-267-763-10	Sequence 10, Appl
32	206.2	20.2	579	9	US-09-454-279-7	Sequence 7, Appl
33	206.2	20.2	579	16	US-10-624-061-7	Sequence 7, Appl
34	204.6	20.1	524	17	US-10-767-795-5016	Sequence 5016, Ap
35	203.8	20.0	1066	13	US-10-425-114-1332	Sequence 1332, Ap
36	201	19.7	1469	17	US-10-437-963-702	Sequence 702, App
37	189.2	18.5	676	17	US-10-767-701-1370	Sequence 1370, Ap
38	173	17.0	622	17	US-10-021-323-4913	Sequence 4913, Ap
39	163.6	16.0	580	17	US-10-021-323-12291	Sequence 12291, A
40	160	15.7	880	13	US-10-425-114-3698	Sequence 3698, Ap
41	152.6	15.0	352	13	US-10-424-599-66073	Sequence 66073, A
42	143.6	14.1	542	17	US-10-021-323-12382	Sequence 12382, A
43	134.6	13.2	2154	17	US-10-437-963-83557	Sequence 83557, A
44	132	12.9	177	13	US-10-424-599-101720	Sequence 101720,
45	129.2	12.7	600	17	US-10-767-795-5596	Sequence 5596, Ap

## ALIGNMENTS

### RESULT 1

US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Farnco, S. Carl  
; APPLICANT: Farnco, Omilayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454, 279  
; EARLIER FILING DATE: 1999-12-03  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Query Match	100.0%	Score 1020;	DB 9;	Length 1020;
Best Local Similarity	100.0%	Pred. No. 2.7e-264;	Mismatches 0;	Indels 0;
Matches 1020;	Conservative 0;			
QY	1	GCACGAGAGCTTAATTCAGAAATGCCCAAGCAGCAGCAAAAGTCTGATTCGGAA	60	
Db	1	GCACGAGAGCTTAATTCAGAAATGCCCAAGCAGCAGCAAAAGTCTGATTCGGAA	60	
QY	61	TGTGTACAGCTAGCTTCTGTGAGAGAGATTGGTAGGCAAGGATACCATTTA	120	
Db	61	TGTGTACAGCTAGCTTCTGTGAGAGAGATTGGTAGGCAAGGATACCATTTA	120	
QY	121	TGGTCTTACATTGACAGAGCAAGTTCCTAGCAATTCACACCTATGATGAAAGTATGC	180	
Db	121	TGGTCTTACATTGACAGAGCAAGTTCCTAGCAATTCACACCTATGATGAAAGTATGC	180	

181 TCAATCCAGGTTTTTGTGGCTCATTTGGTGAATTTGTTGTAAAGATACAGAGGCCAT 240  
181 TCAATCCAGGTTTTTGTGGCTCATTTGGTGAATTTGTTGTAAAGATACAGAGGCCAT 240  
241 TCAAGCTAGGCTGGAGATACAAACCCCTGAAGAAACGGCTTCTCCAGAAATTT 300  
241 TCAAGCTAGGCTGGAGATACAAACCCCTGAAGAAACGGCTTCTCCAGAAATTT 300  
301 ACCACCATCAATTTGTGCCATTTACTCTTCAACAGATTTTTCATCTCTGCTGCTTC 360  
301 ACCACCATCAATTTGTGCCATTTACTCTTCAACAGATTTTTCATCTCTGCTGCTTC 360  
361 AATTACATAAACAGTCCATCTCGAAATATGTAATTTCAAGAGATTTACTTCCATTTGTC 420  
361 AATTACATAAACAGTCCATCTCGAAATATGTAATTTCAAGAGATTTACTTCCATTTGTC 420  
421 TACTTGGGTGATGATGGCAATATGCGCAAACTGAGCTAATGACCTTTTCAATTTATGCA 480  
421 TACTTGGGTGATGATGGCAATATGCGCAAACTGAGCTAATGACCTTTTCAATTTATGCA 480  
481 GTCCATCTPAGAGGATTCATATGGAAGATTTGTTGATGAGGTGAATTTAGGGATGC 540  
481 GTCCATCTPAGAGGATTCATATGGAAGATTTGTTGATGAGGTGAATTTAGGGATGC 540  
541 TCCTCAGACTAGAGCTTTAATTCGAGTAAAGATTAAGAGGATTCATGAAATTTGTT 600  
541 TCCTCAGACTAGAGCTTTAATTCGAGTAAAGATTAAGAGGATTCATGAAATTTGTT 600  
601 GACATTTTACAAGCGTTGAAGAGCGGTGAGGAAGAGATTTGAAGAGCGTGTGGTGT 660  
601 GACATTTTACAAGCGTTGAAGAGCGGTGAGGAAGAGATTTGAAGAGCGTGTGGTGT 660  
661 TGGCAGGAGTCAATTTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 720  
661 TGGCAGGAGTCAATTTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 720  
721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
781 TGAGTACCTCTGGCGGCTAGACTGAAGGATTAATGATGATGATGATGATGATGATGAT 840  
781 TGAGTACCTCTGGCGGCTAGACTGAAGGATTAATGATGATGATGATGATGATGATGAT 840  
841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
901 CAATCTCATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
901 CAATCTCATGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
961 TTTTCATCGCTTTTATAGGAAAGGATTTATATATATATATATATATATATATATATAT 1020  
961 TTTTCATCGCTTTTATAGGAAAGGATTTATATATATATATATATATATATATATATAT 1020

## RESULT 2

US-10-624-061-15  
; Sequence 15, Application US/10624061  
; Publication No. US20040019929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845

; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 15  
; LENGTH: 1020

; TYPE: DNA  
; ORGANISM: Glycine max

US-10-624-061-15

Query Match 100.0%; Score 1020; DB 16; Length 1020;

Best Local Similarity 100.0%; Pred. No. 2,7e-264;

Matches 1020; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCACGAGAGCTTAATTTCAAGAATGGCCAAAGCAGCAGACAAAGCTCTGATTTCTGGGAA 60  
Db 1 GCACGAGAGCTTAATTTCAAGAATGGCCAAAGCAGCAGACAAAGCTCTGATTTCTGGGAA 60  
Qy 61 TGTGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCTTTA 120  
Db 61 TGTGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCTTTA 120  
Qy 121 TGGTCTCATTTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGC 180  
Db 121 TGGTCTCATTTGAGAGAGCCCAAGTTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGC 180  
Qy 181 TCAATCCAGGTTTTTTTGTGGCTCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCAT 240  
Db 181 TCAATCCAGGTTTTTTTGTGGCTCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCAT 240  
Qy 241 TCAAGCTAAGGCTGGAAAGATACAAAACCCCTGAAGAAACGGCTTCTTCCAGAAATTT 300  
Db 241 TCAAGCTAAGGCTGGAAAGATACAAAACCCCTGAAGAAACGGCTTCTTCCAGAAATTT 300  
Qy 301 ACCACCATCAATTTGTGCCATTTACTCTTCAAAACAGTTTTTGCATCTCTGGTCTGCTTC 360  
Db 301 ACCACCATCAATTTGTGCCATTTACTCTTCAAAACAGTTTTTGCATCTCTGGTCTGCTTC 360  
Qy 361 AATTAAATAAACAGTCCATCTGGAAATATGTAATTTCAAGAGTACTTCCATTTGCTTC 420  
Db 361 AATTAAATAAACAGTCCATCTGGAAATATGTAATTTCAAGAGTACTTCCATTTGCTTC 420  
Qy 421 TACTTGGGTGATGATGGCAATATGCGCAAACTGAGCTAATGACCTTTTCAATTTATGCA 480  
Db 421 TACTTGGGTGATGATGGCAATATGCGCAAACTGAGCTAATGACCTTTTCAATTTATGCA 480  
Qy 481 GTCCATCTPAGAGGATTCATATGGAAGATTTGTTGATGAGGTGAATTTAGGGATGC 540  
Db 481 GTCCATCTPAGAGGATTCATATGGAAGATTTGTTGATGAGGTGAATTTAGGGATGC 540  
Qy 541 TCCTCAGACTAGAGCTTTAATTCGAGTAAAGATTAAGAGGATTCATGAAATTTGTT 600  
Db 541 TCCTCAGACTAGAGCTTTAATTCGAGTAAAGATTAAGAGGATTCATGAAATTTGTT 600  
Qy 601 GACATTTTACAAGCGTTGAAGAGCGGTGAGGAAGAGATTTGAAGAGCGTGTGGTGT 660  
Db 601 GACATTTTACAAGCGTTGAAGAGCGGTGAGGAAGAGATTTGAAGAGCGTGTGGTGT 660  
Qy 661 TGGCAGGAGTCAATTTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Db 661 TGGCAGGAGTCAATTTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 720  
Qy 721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
Db 721 ATCAGTGGCTTCTAGCTTTGACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCA 780  
Qy 781 TGAGTACCTCTGGCGGCTAGACTGAAGGATTAATGATGATGATGATGATGATGATGAT 840  
Db 781 TGAGTACCTCTGGCGGCTAGACTGAAGGATTAATGATGATGATGATGATGATGATGAT 840  
Qy 841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
Db 841 TGAAGATGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900



QY 901 CAACTTCATGTAGCTGTTGATTCAGAAATATTAATTTGTAGCTATAGATTGTTACCTTTAT 960  
DB 901 CAACTTCATGTAGCTGTTGATTCAGAAATATTAATTTGTAGCTATAGATTGTTACCTTTAT 960  
QY 961 TTTTCATCGCTTTATTAGGAAAGCATTATATATCATGATCTTCAAAAAAATAAAAAA 1020  
DB 961 TTTTCATCGCTTTATTAGGAAAGCATTATATATCATGATCTTCAAAAAAATAAAAAA 1020  
RESULT 3  
US-10-425-114-7639  
; Sequence 7639, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7639  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700677130\_FLI  
US-10-425-114-7639

Query Match 96.6%; Score 985; DB 13; Length 1079;  
Best Local Similarity 100.0%; Pred. No. 7.9e-255;  
Matches 985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 AGCTTAATTTCAAGATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 67  
DB 95 AGCTTAATTTCAAGATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 154  
QY 68 AGCTTAGCTTCTGTAGAGAGATTGTTAGGCAAGAGATACCATCAATTTATGTTCTC 127  
DB 155 AGCTTAGCTTCTGTAGAGAGATTGTTAGGCAAGAGATACCATCAATTTATGTTCTC 214  
QY 128 ATTGAGAGAGCAAGTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAATC 187  
DB 215 ATTGAGAGAGCAAGTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAATC 274  
QY 188 CAGGTTTTTTGTGCTCATTTGGTGAATTTGTTGTTAAGATACAGAGGCCATTCAAGCT 247  
DB 275 CAGGTTTTTTGTGCTCATTTGGTGAATTTGTTGTTAAGATACAGAGGCCATTCAAGCT 334  
QY 248 AAGCTGGAAGATACAAACCCCTGAAGAAACCCCTTCTCCAGAAATTTACCAACA 307  
DB 335 AAGCTGGAAGATACAAACCCCTGAAGAAACCCCTTCTCCAGAAATTTACCAACA 394  
QY 308 TCAATTGTGCTCATTTCTCTTCAAAACAGTTTTTGTGCTGCTGCTGCTCAATTAAC 367  
DB 395 TCAATTGTGCTCATTTCTCTTCAAAACAGTTTTTGTGCTGCTGCTGCTCAATTAAC 454  
QY 368 ATAAACAGTCCATCTCGAAAATGATTTTCAAGAGATTAATTCATTTGCTTACTTCG 427  
DB 455 ATAAACAGTCCATCTCGAAAATGATTTTCAAGAGATTAATTCATTTGCTTACTTCG 514  
QY 428 GGTGATGATGCAACTATGCGCAACCTGACCTTAATGACCTTTTCATTTATGAGTCCATC 487  
DB 515 GGTGATGATGCAACTATGCGCAACCTGACCTTAATGACCTTTTCATTTATGAGTCCATC 574  
QY 488 TCTAGAGAGATTCATCTATGAAAGTTTGTAGCTGAGGTGAAATTCAGGGATGCTCTCTCAA 547  
DB 575 TCTAGAGAGATTCATCTATGAAAGTTTGTAGCTGAGGTGAAATTCAGGGATGCTCTCTCAA 634

QY 548 GACTACGAGCTTTAATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGTGACATTT 607  
DB 635 GACTACGAGCTTTAATTCGAGCTAAGGATAAAGAGGATTGATGAAATTTGTGACATTT 694  
QY 608 ACAAGCGTTGAAGAGACGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTTTGGGCGAG 667  
DB 695 ACAAGCGTTGAAGAGACGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTTTGGGCGAG 754  
QY 668 GAAGTGAATCTTAAACAGTATGATGACATGACATGAAACCGTAAATTTGATCCATCAGTG 727  
DB 755 GAAGTGAATCTTAAACAGTATGATGACATGACATGAAACCGTAAATTTGATCCATCAGTG 814  
QY 728 GCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTAC 787  
DB 815 GCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTAC 874  
QY 788 CTCTTGGCGGCTAGACTGAAGGCATTACATGCAATGCAATTTAGAAATGGAAGA 847  
DB 875 CTCTTGGCGGCTAGACTGAAGGCATTACATGCAATGCAATTTAGAAATGGAAGA 934  
QY 848 TGAATATGATTTGTTGTTTCAATGATTAAGCTCTTAAGTGTATCCCTTTATTCCTCACTTC 907  
DB 935 TGAATATGATTTGTTGTTTCAATGATTAAGCTCTTAAGTGTATCCCTTTATTCCTCACTTC 994  
QY 908 ATGTAGCTGTTGATTCAGAAATTAATTTGTAGCTATAGATTGTTACCTTTATTTTCATC 967  
DB 995 ATGTAGCTGTTGATTCAGAAATTAATTTGTAGCTATAGATTGTTACCTTTATTTTCATC 1054  
QY 968 GCGTTTTATTAGGAAAGGCAATTATT 992  
DB 1055 GCGTTTTATTAGGAAAGGCAATTATT 1079

RESULT 4  
US-10-424-599-93036  
; Sequence 93036, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 93036  
; LENGTH: 1653  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55025C.1  
US-10-424-599-93036

Query Match 92.2%; Score 940.4; DB 13; Length 1653;  
Best Local Similarity 96.6%; Pred. No. 1.1e-242;  
Matches 972; Conservative 0; Mismatches 31; Indels 3; Gaps 1;  
QY 8 AGCTTAATTTCAAGAAATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 67  
DB 432 AGCTTAATTTCAAGAAATGCGCAAGCAGCAGCAACAAAGTCTGATTTCTGGGAATGTGTAC 491  
QY 68 AGCTTAGCTTCTGTAGAGAGATTGTTAGGCAAGAGGATACCATCATTTATGCTTC 127  
DB 492 AGCTTAGCTTCTGTAGAGAGATTGTTAGGCAAGAGGATACCATCATTTATGCTTC 551  
QY 128 ATTGAGAGAGCAAGTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAATC 187  
DB 552 ATTGAGAGAGCAAGTCCCTAGCAATTTCTCACACCTATGATGAAAGATGCTCAATC 611

188 CAGGGTTTTTGGGCTCAATGGTGGAAATTTGTTTAAAGATACAGAGCCCAATCAAGCT 247  
189 |  
190 |  
191 |  
192 |  
193 |  
194 |  
195 |  
196 |  
197 |  
198 |  
199 |  
200 |  
201 |  
202 |  
203 |  
204 |  
205 |  
206 |  
207 |  
208 |  
209 |  
210 |  
211 |  
212 |  
213 |  
214 |  
215 |  
216 |  
217 |  
218 |  
219 |  
220 |  
221 |  
222 |  
223 |  
224 |  
225 |  
226 |  
227 |  
228 |  
229 |  
230 |  
231 |  
232 |  
233 |  
234 |  
235 |  
236 |  
237 |  
238 |  
239 |  
240 |  
241 |  
242 |  
243 |  
244 |  
245 |  
246 |  
247 |  
248 |  
249 |  
250 |  
251 |  
252 |  
253 |  
254 |  
255 |  
256 |  
257 |  
258 |  
259 |  
260 |  
261 |  
262 |  
263 |  
264 |  
265 |  
266 |  
267 |  
268 |  
269 |  
270 |  
271 |  
272 |  
273 |  
274 |  
275 |  
276 |  
277 |  
278 |  
279 |  
280 |  
281 |  
282 |  
283 |  
284 |  
285 |  
286 |  
287 |  
288 |  
289 |  
290 |  
291 |  
292 |  
293 |  
294 |  
295 |  
296 |  
297 |  
298 |  
299 |  
300 |  
301 |  
302 |  
303 |  
304 |  
305 |  
306 |  
307 |  
308 |  
309 |  
310 |  
311 |  
312 |  
313 |  
314 |  
315 |  
316 |  
317 |  
318 |  
319 |  
320 |  
321 |  
322 |  
323 |  
324 |  
325 |  
326 |  
327 |  
328 |  
329 |  
330 |  
331 |  
332 |  
333 |  
334 |  
335 |  
336 |  
337 |  
338 |  
339 |  
340 |  
341 |  
342 |  
343 |  
344 |  
345 |  
346 |  
347 |  
348 |  
349 |  
350 |  
351 |  
352 |  
353 |  
354 |  
355 |  
356 |  
357 |  
358 |  
359 |  
360 |  
361 |  
362 |  
363 |  
364 |  
365 |  
366 |  
367 |  
368 |  
369 |  
370 |  
371 |  
372 |  
373 |  
374 |  
375 |  
376 |  
377 |  
378 |  
379 |  
380 |  
381 |  
382 |  
383 |  
384 |  
385 |  
386 |  
387 |  
388 |  
389 |  
390 |  
391 |  
392 |  
393 |  
394 |  
395 |  
396 |  
397 |  
398 |  
399 |  
400 |  
401 |  
402 |  
403 |  
404 |  
405 |  
406 |  
407 |  
408 |  
409 |  
410 |  
411 |  
412 |  
413 |  
414 |  
415 |  
416 |  
417 |  
418 |  
419 |  
420 |  
421 |  
422 |  
423 |  
424 |  
425 |  
426 |  
427 |  
428 |  
429 |  
430 |  
431 |  
432 |  
433 |  
434 |  
435 |  
436 |  
437 |  
438 |  
439 |  
440 |  
441 |  
442 |  
443 |  
444 |  
445 |  
446 |  
447 |  
448 |  
449 |  
450 |  
451 |  
452 |  
453 |  
454 |  
455 |  
456 |  
457 |  
458 |  
459 |  
460 |  
461 |  
462 |  
463 |  
464 |  
465 |  
466 |  
467 |  
468 |  
469 |  
470 |  
471 |  
472 |  
473 |  
474 |  
475 |  
476 |  
477 |  
478 |  
479 |  
480 |  
481 |  
482 |  
483 |  
484 |  
485 |  
486 |  
487 |  
488 |  
489 |  
490 |  
491 |  
492 |  
493 |  
494 |  
495 |  
496 |  
497 |  
498 |  
499 |  
500 |  
501 |  
502 |  
503 |  
504 |  
505 |  
506 |  
507 |  
508 |  
509 |  
510 |  
511 |  
512 |  
513 |  
514 |  
515 |  
516 |  
517 |  
518 |  
519 |  
520 |  
521 |  
522 |  
523 |  
524 |  
525 |  
526 |  
527 |  
528 |  
529 |  
530 |  
531 |  
532 |  
533 |  
534 |  
535 |  
536 |  
537 |  
538 |  
539 |  
540 |  
541 |  
542 |  
543 |  
544 |  
545 |  
546 |  
547 |  
548 |  
549 |  
550 |  
551 |  
552 |  
553 |  
554 |  
555 |  
556 |  
557 |  
558 |  
559 |  
560 |  
561 |  
562 |  
563 |  
564 |  
565 |  
566 |  
567 |  
568 |  
569 |  
570 |  
571 |  
572 |  
573 |  
574 |  
575 |  
576 |  
577 |  
578 |  
579 |  
580 |  
581 |  
582 |  
583 |  
584 |  
585 |  
586 |  
587 |  
588 |  
589 |  
590 |  
591 |  
592 |  
593 |  
594 |  
595 |  
596 |  
597 |  
598 |  
599 |  
600 |  
601 |  
602 |  
603 |  
604 |  
605 |  
606 |  
607 |  
608 |  
609 |  
610 |  
611 |  
612 |  
613 |  
614 |  
615 |  
616 |  
617 |  
618 |  
619 |  
620 |  
621 |  
622 |  
623 |  
624 |  
625 |  
626 |  
627 |  
628 |  
629 |  
630 |  
631 |  
632 |  
633 |  
634 |  
635 |  
636 |  
637 |  
638 |  
639 |  
640 |  
641 |  
642 |  
643 |  
644 |  
645 |  
646 |  
647 |  
648 |  
649 |  
650 |  
651 |  
652 |  
653 |  
654 |  
655 |  
656 |  
657 |  
658 |  
659 |  
660 |  
661 |  
662 |  
663 |  
664 |  
665 |  
666 |  
667 |  
668 |  
669 |  
670 |  
671 |  
672 |  
673 |  
674 |  
675 |  
676 |  
677 |  
678 |  
679 |  
680 |  
681 |  
682 |  
683 |  
684 |  
685 |  
686 |  
687 |  
688 |  
689 |  
690 |  
691 |  
692 |  
693 |  
694 |  
695 |  
696 |  
697 |  
698 |  
699 |  
700 |  
701 |  
702 |  
703 |  
704 |  
705 |  
706 |  
707 |  
708 |  
709 |  
710 |  
711 |  
712 |  
713 |  
714 |  
715 |  
716 |  
717 |  
718 |  
719 |  
720 |  
721 |  
722 |  
723 |  
724 |  
725 |  
726 |  
727 |  
728 |  
729 |  
730 |  
731 |  
732 |  
733 |  
734 |  
735 |  
736 |  
737 |  
738 |  
739 |  
740 |  
741 |  
742 |  
743 |  
744 |  
745 |  
746 |  
747 |  
748 |  
749 |  
750 |  
751 |  
752 |  
753 |  
754 |  
755 |  
756 |  
757 |  
758 |  
759 |  
760 |  
761 |  
762 |  
763 |  
764 |  
765 |  
766 |  
767 |  
768 |  
769 |  
770 |  
771 |  
772 |  
773 |  
774 |  
775 |  
776 |  
777 |  
778 |  
779 |  
780 |  
781 |  
782 |  
783 |  
784 |  
785 |  
786 |  
787 |  
788 |  
789 |  
790 |  
791 |  
792 |  
793 |  
794 |  
795 |  
796 |  
797 |  
798 |  
799 |  
800 |  
801 |  
802 |  
803 |  
804 |  
805 |  
806 |  
807 |  
808 |  
809 |  
810 |  
811 |  
812 |  
813 |  
814 |  
815 |  
816 |  
817 |  
818 |  
819 |  
820 |  
821 |  
822 |  
823 |  
824 |  
825 |  
826 |  
827 |  
828 |  
829 |  
830 |  
831 |  
832 |  
833 |  
834 |  
835 |  
836 |  
837 |  
838 |  
839 |  
840 |  
841 |  
842 |  
843 |  
844 |  
845 |  
846 |  
847 |  
848 |  
849 |  
850 |  
851 |  
852 |  
853 |  
854 |  
855 |  
856 |  
857 |  
858 |  
859 |  
860 |  
861 |  
862 |  
863 |  
864 |  
865 |  
866 |  
867 |  
868 |  
869 |  
870 |  
871 |  
872 |  
873 |  
874 |  
875 |  
876 |  
877 |  
878 |  
879 |  
880 |  
881 |  
882 |  
883 |  
884 |  
885 |  
886 |  
887 |  
888 |  
889 |  
890 |  
891 |  
892 |  
893 |  
894 |  
895 |  
896 |  
897 |  
898 |  
899 |  
900 |  
901 |  
902 |  
903 |  
904 |  
905 |  
906 |  
907 |  
908 |  
909 |  
910 |  
911 |  
912 |  
913 |  
914 |  
915 |  
916 |  
917 |  
918 |  
919 |  
920 |  
921 |  
922 |  
923 |  
924 |  
925 |  
926 |  
927 |  
928 |  
929 |  
930 |  
931 |  
932 |  
933 |  
934 |  
935 |  
936 |  
937 |  
938 |  
939 |  
940 |  
941 |  
942 |  
943 |  
944 |  
945 |  
946 |  
947 |  
948 |  
949 |  
950 |  
951 |  
952 |  
953 |  
954 |  
955 |  
956 |  
957 |  
958 |  
959 |  
960 |  
961 |  
962 |  
963 |  
964 |  
965 |  
966 |  
967 |  
968 |  
969 |  
970 |  
971 |  
972 |  
973 |  
974 |  
975 |  
976 |  
977 |  
978 |  
979 |  
980 |  
981 |  
982 |  
983 |  
984 |  
985 |  
986 |  
987 |  
988 |  
989 |  
990 |  
991 |  
992 |  
993 |  
994 |  
995 |  
996 |  
997 |  
998 |  
999 |  
1000 |

RESULT 5  
US-10-425-114-15059  
; Sequence 15059, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 15059  
; LENGTH: 992  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3050-017-B9\_FLI  
US-10-425-114-15059

Query Match 67.6%; Score 689.4; DB 13; Length 992;  
Best Local Similarity 92.6%; Pred. No. 4e-175;  
Matches 747; Conservative 0; Mismatches 56; Indels 4; Gaps 2;

QY 208 GGTGGAAATTTGTTTAAAGAAATACAGAGCCCAATCAAGCTAGAGCTGGGAAGATACAAAAA 267  
DB 1 GGTAGAAATTTGTTTAAAGAAATACAGAGCCCAATCAAGCTAGAGCTGGGAAGATACAAAAA 60  
QY 268 CCCTGAAGAAACGCGCTTCTCCAGAAAAATTTACCAATCAATTTGTCCTTCTTACTC 327  
DB 61 CCCTGAAGAAACGCGCTTCTCCAGAAAAATTTACCAATCAATTTGTCCTTCTTACTC 120  
QY 328 CTTCAACACAGTTTTTCATCTGCTGCTTCAATTTAAACATAAACAAGTCCATCTGGA 387  
DB 121 CTTCAACACAGTTTTTCATCTGCTGCTTCAATTTAAACATAAACAAGTCCATCTGGA 180  
QY 388 AATGTAATTTCAAGAGTTTACTTCCATTTGCTTCTTCTCGGTTGATGAGGCAATCTATGC 447  
DB 181 AATGTAATTTCAAGAGTTTACTTCCATTTGCTTCTTCTCGGTTGATGAGGCAATCTATGC 240  
QY 448 GCAATCTGAGCTTAATGACCTTTTCAATTTGCTGCTTCTTCTAGAGGATTCACCTATG 507  
DB 241 ACAATCTGAGCTTAATGACCTTTTCAATTTGCTGCTTCTTCTAGAGGATTCACCTATG 300  
QY 508 AAGTTTGTAGCTGAGGTGAATTTCAAGGATGCTCTCAAGACTACGAGCTTTAATTCG 567  
DB 301 AAGTTTGTAGCTGAGGTGAATTTCAAGGATGCTCTCAAGACTACGAGCTTTAATTCG 360  
QY 568 AGCTTAAGGATTAAGAGGATTTGATGAAATTTGTTGACATTTACAGGCTTGAAGAGACGCT 627  
DB 361 AGCTTAAGGATTAAGAGGATTTGATGAAATTTGTTGACATTTACAGGCTTGAAGAGACGCT 420  
QY 628 GAGGACAGAGTTTGAAGAGAGGCTGTGTTGTTGGGAGGAGTGAATCTTAAACAGTGA 687  
DB 421 GAGGACAGAGTTTGAAGAGAGGCTGTGTTGTTGGGAGGAGTGAATCTTAAACAGTGA 480  
QY 688 TGCAATGACAAATGAAGAACCGTAAATTTGATCCATCGATGGCTTCTAGCTTGTACAAAA 747  
DB 481 GGACGATGATGAGGAAAAACCAATTAAGTTTGTATCCATCAGTGGCTTCTAGCTTGTACAAAA 540  
QY 748 TTGGGTGATACCTCTCCAGAGGAGGTTCAAGTTGAGTACCTCTTGGCGCGCTTACAGCTG 807  
DB 541 TTGGGTGATACCTCTCCAGAGGAGGTTCAAGTTGAGTACCTCTTGGCGCGCTTACAGCTG 600  
QY 808 AAGGCATTACAATGACGTTAGAAATTTAGAGAAATGGAAGATGAATATGATGTTGTTGTTTC 867  
DB 601 AAGGCATTACAATGACGTTAGAAATTTAGAGAAATGGAAGATGAATATGATGTTGTTGTTTC 660  
QY 868 AATGATTAAGCTTTAAGTATCCCTTTTATTTGCAACTTCAATCTAGCTTGTGATGAGAA 927  
DB 661 AATGATTAAGCTTTAAGTATCCCTTTTATTTGCAACTTCAATCTAGCTTGTGATGAGAA 720  
QY 928 ATATT-ATTTGTAGCTATAGATTTTACCTTTTATTTTCTCGGCTTTATTT-AGGAAAA 983  
DB 721 ATATTAAATTTGACCTTTAGATTTGTTACTTTTATTTTCTCGGCTTTATTTAGAGAAAA 780  
QY 984 GGCATTATATATCATGATCTTCAAAA 1010  
DB 781 AAAAGGATTAGATCATGAATTTCAATA 807

RESULT 6  
US-09-454-279-5  
; Sequence 5, Application US/09454279



```
Publication No. US20030077687A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jorn
APPLICANT: Boyes, Douglas
APPLICANT: Davis, Keith
APPLICANT: Hamilton, Carol
APPLICANT: Hoffman, Neil
APPLICANT: Kloti, Andreas
APPLICANT: Woessner, Jeffrey
APPLICANT: Zayed, Adel
APPLICANT: Ascenzi, Robert
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
FILE REFERENCE: 2035DIV1
CURRENT APPLICATION NUMBER: US/10/267,763
CURRENT FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: 09/610,040
PRIOR FILING DATE: 2000-07-05
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO. 3
LENGTH: 1006
TYPE: DNA
ORGANISM: Arabidopsis
US-10-267-763-3

Query Match 32.1%; Score 327.2; DB 15; Length 1006;
Best Local Similarity 64.2%; Pred. No. 2e-77;
Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGCCAAAGCAGCAGACAAAGTCTGATCTGGGAATGTGACACGCTAGCTTCTGT 81
DB 69 AAGAGCTTCGAATCGGATTCGGTGTCTGTTCCAAATGACTGAGCTTGAATTAAT 128
QY 82 GAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTCANTGAGAGACCAA 141
DB 129 CAGAGAATCGTTGATTAGGCAAGACACCATCGTCTTACGCTTATCGAGAGAGCTAA 188
QY 142 GTTCCCTAGCAATCTCACACCTATGATGAAAGTATGCTCAAAATCCAGGGTTTTGTGG 201
DB 189 GTTCCCTAGCAATCTCACACCTATGATGAAAGTATGCTCAAAATCCAGGGTTTTGTGG 248
QY 202 CTCATTTGGGAATTTGTTAGAAATACAGAGCCATTCAGCTTACGCTTGAAGATA 261
DB 249 TTCTCTCACTGAGTATTCCTGTCAGAGACAAATCATCCAGCTTAAGGTAGGAGATA 308
QY 262 CAAAAACCCCTGAAGAAAGCCCTTCTCCAGAAATTTACCACCATCAATTTGTGCCATC 321
DB 309 TGAATACCCGAGAGAGATCTTCTCTTCTTGAACATTTCTCACTCGGTTTTCCTAC 368
QY 322 TTACTCCTTCAACAGTTTTCATCCTGGTGTGCTTCAATTAACAACAGTCCAT 381
DB 369 GCACAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTTAAATTAACAAACAAT 428
QY 382 CTGAAAAATGATTTTCAAGAGTACTTCCATTGCTTGTCTCTGGGTGATGATGGCAA 441
DB 429 CTGGGATTTTACTTTAAGAAATGCTTCTCTTGTGTTGTTCAACCTGGGATGAGGCAA 488
QY 442 CTATGCGCAAACTGACGATTAAGACCTTTCAATTTATGAGTGCATCTCTAGAGGATCA 501
DB 489 CTATCACTCAACTGCTGATGATCTCGCTGTTTACAAAGTCTTTCGAGAGGATTTCA 548
QY 502 CTATGAGAAATTTGATGAGTGAATTCAGGATGCTCTCAAGACTAGAGCTTTT 561
DB 549 CTACGGTAAATTTGATGAGTGAATTCAGAGATGCTTCCAGAGATTAAGAGCTTGC 608
QY 562 AATTCGAGCTAAGGATTAAGAGATTTGATGAAATTTGATGATTAAGAGCTTTGAAGA 621
DB 609 GATTCGGCTCAGGATAGAGGCTTTGATGAGGCTTTGATGAGGCTTTGAGAAAGTGAAGA 668
QY 622 GACGTTGAGGAGAGGTTCAAGAGAGGCTTGGTGTGTTGGGCGAGGAGTCAATCTTAA 681
```

1669 AATGGTTTAAAGAGAGAGTGCAGAAAGAAAGCAGAAAGCGTTTGGACAGAAAGTAAATTTCAA 728

1682 -----CAGTGATGACAATGACAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAG 735

1729 CTCGGCTATGGCGATGAGAGTAAGAAAGATATAAGTGGATCCATTCGTTCCCTCTCG 788

1736 CTTGTACAAAAATTTGGTGATACCTCTCCACCAAGGAGGTTTCAAGTTGAGTACCTCTTGG 795

1789 CACTACGGGGAATGGCTTATCCCTCTCACTAAGCTGTTGAGGTTGAGTATCTTCTACG 848

1796 CCGCTAGACTGAA 809

1849 TCGCTCGATTGAA 862

RESULT 9

US-10-267-763-9/c

Sequence 9, Application US/10267763

Publication No. US20030077687A1

GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn

APPLICANT: Boyes, Douglas

APPLICANT: Davis, Keith

APPLICANT: Hamilton, Carol

APPLICANT: Hoffman, Neil

APPLICANT: Kloti, Andreas

APPLICANT: Woessner, Jeffrey

APPLICANT: Zayed, Adel

APPLICANT: Ascenzi, Robert

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE

MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE

TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE

FILE REFERENCE: 2035DIV1

CURRENT APPLICATION NUMBER: US/10/267,763

CURRENT FILING DATE: 2002-10-09

PRIOR APPLICATION NUMBER: 09/610,040

PRIOR FILING DATE: 2000-07-05

NUMBER OF SEQ ID NOS: 11

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO. 9

LENGTH: 1006

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-10-267-763-9

Query Match 32.1%; Score 327.2; DB 15; Length 1006;

Best Local Similarity 64.2%; Pred. No. 2e-77;

Matches 510; Conservative 0; Mismatches 278; Indels 6; Gaps 1;

QY 22 AATGCCAAAGCAGCAGACAAAGTCTGATCTGGGAATGTGACACGCTAGCTTCTGT 81

DB 938 AAGAGCTTCGAATCGGATTCGGTGTCTGTTCCAAATGACTGAGTCTTGAATTAAT 879

QY 82 GAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTCATTTGAGAGACCAA 141

DB 878 CAGAGAATCGTTGATTAGGCAAGACACCATCGTCTTCAAGCTTGTATCGAGAGACTAA 819

QY 142 GTTCCCTAGCAATTTCTCACACCTATGATGAAAGTATGCTCAAAATCCAGGGTTTTGTGG 201

DB 818 GTTTCCTCAATTTCTCTGCTTTCGAGGATCTGTTCTAGATTTCTGGAAGTTCTC 759

QY 202 CTCATTCGTTGGAATTTGTTTAAAGATAACAGAGGCCATTCAGAGCTTAAGCTGGAAGATA 261

DB 758 TTCTCTCACTGAGTATTTCTGTCAGAGAGACAGAAATCATCCAGACTTAAGGTAGGAAGATA 699

QY 262 CAAAAACCCCTGAAGAAAGCCCTTCTCCAGAAATTTACCACCATCAATTTGTGCCATC 321

DB 698 TGAATACCCGAGAGAGATCTTCTTCTCTTGAACATTTCTCACTCGGTTTTCCTAC 639

QY 322 TTACTCCTTCAACAGTTTTCATCCTGGTGTGCTTCAATTAACAACAGTCCAT 381

DB 638 GCACAAATATCCATCGGCTTTGACCCCTAGGCTCTATCTGTTAAATTAACAAACAAT 579

382	CTGGAAATGTATTTCAAGAGATTACTTCATTCGTTGCTACTCTCGGGTGATGATGGCAA	441
Qy		
578	CTGGGATATTACTTTAAAGAAATTCGTTCTCTTGTGTTGTCAAACCTGCGATGATGGCAA	519
Db		
442	CTATGGCGAAACTGCAGCTTAATGACCTTTTCATTATTGCAGTCCATCTCTAGAAGGATTCA	501
Qy		
518	CTATCCATCAATGCTGCTAGTAGTGAATCTGCGCTGTTTACAGCTCTTTTGGAGAAGGATTCA	459
Db		
502	CTATGGAAGTTTGTAGCTGAGGTGAATTCAGGGATGCTCCTCAAGACTACGAGCGCTTT	561
Qy		
458	CTACGGTAAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAGCGCTCG	399
Db		
562	AATTCGAGCTTAAGGATAAAGAGGATTTCATGAATTTGTTGACATTTTCAAGCGTTTGARGA	621
Qy		
398	GATTCGCGTCAAGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAGAAAGTAGAAGA	339
Db		
622	GACGTTGAGGAAGAGAGTTTGAAGAAAGCGCTGTGCTGTTTTGGCGAGGAAGTGAATCTTAA	681
Qy		
338	AATGGTTAAGAAGAGAGTGCAGAAGAAAGCAGAAACGTTTGGACAAGAAGTAAATTCAA	279
Db		
682	-----CAGTGATGACAATGACAATGAAGAACCGTTAAATTTGATCCATCAGTGGCTTCTAG	735
Qy		
278	CTCTGGCTATGCGCATGAGATGAAGAAGAAGTATAAAGTGGATCCATTGCTTGGCTCTCG	219
Db		
736	CTTTGTACAAAATTTGGGTGATACCTCTCACCAAGAGGTTTCAGTTGAGTACCTCTTTGG	795
Qy		
218	CATCTACGGGGAATGGCTTATCCCTCTCACCTTAAGCTCGTTGAGGTTGAGTATCTTCTACG	159
Db		
796	CGGCTPAGACTGAA	809
Qy		
158	TGGTCTCGATTGAA	145
Db		

## RESULT 10

```

US-09-938-842A-972
; Sequence 972, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 972
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-972

```

	Query Match	32.0%;	Score 326.2;	DB 9;	Length 798;
	Best Local Similarity	64.2%;	Pred. No. 3.2e-77;		
	Matches 509;	Conservative 0;	Mismatches 278;	Indels 6;	Gaps 1
Qy	22	AATGCCAAAGCAGCAGAGAACAAAGTCCTGATTCGGGAATGTGTACACGCTAGCTTCTGT	81		
Db	6	AAGAGCTTTCGAATCGGATTCGGGTTCGGTGTGCCAATGTACTGAGTCTTGACTTAAT	65		
Qy	82	GAGAGAGGATTTGGTTAGCGCAGAGAGTACCATCATTTATGGTCTCATTCAGAGAGCCAA	141		
Db	66	CHAGAGATCGTTGATTTGCGCAAGAGACACCATCGTCTTCAGTTCGATCGAGAGAGTAA	125		
Qy	142	GTTCCTTAGCAATTCACACATATGATGAAAAGTATGCTCAAAATCCAGGGTTTTTGTGG	201		

[illegible]

RESIT. 11

```

RES001.11
US-09-338-842A-972
  Sequence 972, Application US/0938842A
  Publication No. US20040009476A9
  GENERAL INFORMATION:
  APPLICANT: Harper, Jeff
  APPLICANT: Kreps, Joel
  APPLICANT: Zhu, Xun
  APPLICANT: Zhang, Tong
  TITLE OF INVENTION: STRESS-REGULATED
  TITLE OF INVENTION: SAME, AND METHOD
  FILE REFERENCE: SCRI300-3
  CURRENT APPLICATION NUMBER: US/09/338
  CURRENT FILING DATE: 2001-08-24
  PRIOR APPLICATION NUMBER: US 60/227,8
  PRIOR FILING DATE: 2000-08-24
  PRIOR APPLICATION NUMBER: US 60/264,6
  PRIOR FILING DATE: 2001-01-16
  PRIOR APPLICATION NUMBER: US 60/300,1
  PRIOR FILING DATE: 2001-06-22
  NUMBER OF SEQ ID NOS: 5379
  SEQ ID NO 972
  LENGTH: 798
  TYPE: DNA
  ORGANISM: Arabidopsis thaliana

```

APPLICANT: ZHU, Tong  
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING





; SEQ ID NO 93956  
; LENGTH: 1016  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_92290C.1  
US-10-437-963-93956

Query Match 25.2%; Score 257.4; DB 17; Length 1016;  
Best Local Similarity 58.4%; Pred. No. 1.3e-58;  
Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;  
Qy 24 TGGCCAAAGCAGCAGACAAAGCTCTGATCTGGATGTGTACACGCTAGCTTCTGTGA 83  
Db 155 TGACCAAGGAGGAGAAACAGAGATAGATCAAGTGAATFACTGACCTTGGACAACATTA 214  
Qy 84 GAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGCTCTCATTTAGAGAGACCAAGT 143  
Db 215 GAACCTCCTCTGGTTAGGCAAGAGACAGCATCATTTATTCAGCCCTTTAGAGAGACAGT 274  
Qy 144 TCCCTAGCAATTCACACCTATGATGAAAAGTATGCTCAATCCAGGGTTTTTGTGGCT 203  
Db 275 TTTGCTCAATGCTGATATATATGATAAAATGCTTTCCATGTGGATGGATTTGATGGCT 334  
Qy 204 CATTTGGTGGAAATTTGTTTAAAGAAATACAGAGGCCATTCAGCTAAGGCTGGAGATACA 263  
Db 335 CTTTGGTTGATTCATGTTAGAGAAACCGAAACCTACATCAACAGGTTGGGAGATACA 394  
Qy 264 AAAACCTGAGAAAACCGCTTCTTCCAGAAAATTTACACCATCAATTTGTGCCATCTT 323  
Db 395 AGAGCCCTGATGAGACCCATCTTCTCCGGAGGATCTGCCGAACCACTTTGCCACCTC 454  
Qy 324 ACTCTTTCAACACGTTTTTGCATCTGCTGCTCAATTAACATAAACAAGTCCATCT 383  
Db 455 TCCAGTATCCAAGGTTTTGCACTTATGCTGATCTATTAATATCAACAAGAGATT 514  
Qy 384 GGAATAATGATTTCAAGAGATTAATTCATTTGCTGCTTCCATTTGCTGCTCGGGTATGATGCAACT 443  
Db 515 GGAATAATGATTTTGTAGTGGCTTCTTCCAAAGATTTAGTGAAGAAGGAGATGATGTAATT 574  
Qy 444 ATGGCAAACTGACGCTTAAGCCTTTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503  
Db 575 ATGGATCCAGTCTCTTTGTGACACGATCTGCTTGCAGGCGCTCTCCAAAAGAAATTCAC 634  
Qy 504 ATGAAAAGTTTGTAGCTGAGGTGAATTCAGGGATGCTCTCAAGACTACAGCCTTTAA 563  
Db 635 ATGTAAGTTTGTGCGAGAGCTAAGTTTCAAGAGTCTCTGAGCTTACATGCTGCGA 694  
Qy 564 TTCGAGCTAAGGATAAAGAGGATGATGAATTTGTTGACATTTACAGGCTTGAAGAGA 623  
Db 695 TAAATAGCACAGACTGCGATCAACTAATGCACTCTCTCACCTATGAACCGTGGAGG 754  
Qy 624 CGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGGTGTGTTGGCGAGGAGTGAATCTTAA 683  
Db 755 CTATGTAACATAGGTTGAGCTAAGCTAAGCTAAGATCTTTGACAGGAGGTTAGGCG 814  
Qy 684 GTGATGACAAATGAACAAT---GAAAACCGTAATTTGATTCATCATGCTGCTTCTAGCTGT 740  
Db 815 CTGAAGACACAGCGCTCCACCAATGTACAAGATAAGGCCCAAGTTTGGTGGCTGAAC 874  
Qy 741 ACAAATAATGGTGATACCTCTCAAGAGGTTTCAAGTTGAGTACCTCTTGGCGGCTC 800  
Db 875 ACAGTACAGGATCATGCGGCTAACCAGGAGGTTCAAGTAGCCCTCTTCTGAGGAGAT 934  
Qy 801 TAGACTGAAGGCATTACAATGCA 823  
Db 935 TGGATTGATTTTACGATTGTA 957

## RESULT 14

US-09-454-279-13  
; Sequence 13, Application US/09454279  
; Publication No. US20020184659A1

; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 1231  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-454-279-13

Query Match 25.2%; Score 257.4; DB 9; Length 1231;  
Best Local Similarity 58.4%; Pred. No. 1.5e-58;  
Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;  
Qy 24 TGGCCAAAGCAGCAGACAAAGTCTGATCTGGAAATGTGTACACGCTAGCTTCTGTGA 83  
Db 239 TGACCAAGGAGGAGAAACAGAGATAGATCAAAAGTGAATACTGACCTTGGACAACATTA 298  
Qy 84 GAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGCTCTCATTTGAGAGACCAAGT 143  
Db 299 GAACCTCCTCTGGTTAGGCAAGAGACAGCATCATTTAGCCTCTTAGAGAGACAGT 358  
Qy 144 TCCCTAGCAATTTCTCACACCTATGATGAAAAGTATGCTCAAAATCCAGGGTTTTGTGGCT 203  
Db 359 TTTGCTACAATCTCATATATATGATAAAATGCTTTCCATGTGGATGGATTTGATGGCT 418  
Qy 204 CATTTGGTGGAAATTTGCTTAAAGATACAGAGGCCATTTACAGCTAAGCTTGAAGATACA 263  
Db 419 CTTTGGTGAATTCATGTTAGAGAAACCGAAAACTACATCAACAGCTTGGAGATACA 478  
Qy 264 AAAACCTTGAAGAAAACGCTTTCTTCCAGAAAATTTACCACTCAATTTGTGCCATCTT 323  
Db 479 AGAGCCCTGATGAGCACCCATCTTTTCCGAGGATCTGCTGAAACCACTGTTGCCACCTC 538  
Qy 324 ACTCCTTCAACAGTTTTTGCATCTGCTGCTGCTTCAATTAACATAAACAAGTCCATCT 383  
Db 539 TCCAGTATCCAAAGGTTTTGCATCTCTTGTGATTTCTTAAATATCAACAAGGAGATT 598  
Qy 384 GGAATAATGATTTCAAGAGTTTACTTCCATTTGCTGCTTCCAGTGTGATGATGCAACT 443  
Db 599 GGAATAATGATTTGATGAGCTTCTTCCAGATAGTGAAGAGAGAGTGAATGTAAT 658  
Qy 444 ATGGCAAACTGACGCTAATGACCTTTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503  
Db 659 ATGGATCCAGTCTCTTTGTGACAGCATCTGCTTGCAGGCGCTCTCCAAAAGAAATTCAC 718  
Qy 504 ATGGAAGTTTGTAGCTGAGGTGAATTTCAAGGATGCTCTCAAGACTACAGGCTTTAA 563  
Db 719 ATGGTAAGTTTGTGCGAGAGGCTAAGTTTCAAGAGTCTCTTGAAGCTTACATGCTCGGA 778  
Qy 564 TTCGAGCTAAGGATAAAGAGGATGATGAATTTTGCACATTTTACAAGCGTTGAAGAGA 623  
Db 779 TAAATAGCAGGACTGCGATCACTCACTCTTCCAGCTTCACTGAAACGGTGGAGCGTG 838  
Qy 624 CGGTGAGGAGAGAGTTGAAAAGAGGCTGTGGTGTGTTGGCGAGGAGTGAATCTTAA 683  
Db 839 CTATTGAACATAGGTGGAGGCTAAGCTTAAAGCTTAAGATCTTTGGACAGAGGTGGATTAGCG 898  
Qy 684 GTGATGACAAATGACAAT---GAAAACCGTAAATTTGATCTCACTGCTTCTAGCTGT 740  
Db 899 CTGAAGACAAACGCGCTCCACCAATGTACAGATAAGGCCAGTTTGGTGGCTGAATGT 958  
Qy 741 ACAAATAATGGTGATACCTCTCAAGAGGAGTTTCAAGTTGAGTACCTCTTGGCGGCTC 800  
Db 959 ACAGCTTACAGGATCATGCGGCTTAACCAAGAGGTTTCAAGTAGCCCTACTTGTGAGGAGAT 1018



QY 801 TAGACTGAAGGCAATTACAATGCA 823  
Db 1019 TGGATTGATTGTTTACGATTGTA 1041

RESULT 15

US-10-624-061-13  
; Sequence 13, Application US/10624061  
; Publication No. US20040019929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Ramodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 1231  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-10-624-061-13

Query Match 25.2%; Score 257.4; DB 16; Length 1231;  
Best Local Similarity 58.4%; Pred. No. 1.5e-58;  
Matches 469; Conservative 0; Mismatches 331; Indels 3; Gaps 1;

QY 24 TGGCCAAAGCAGCAGAAAGTCCTGATTCGGGAATGTGTACACGCTAGCTTCTGTGA 83  
Db 239 TGACCAAGGAGGAGAAACAGAGATAGATCAAAAGTGAATCTGACCTTGGACAACTTA 298

QY 84 GACAGGATTGGTTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGAGCAAGT 143  
Db 299 GAACCTCTTGGTTAGGCAAGAGAGAGATATATTCAGCCTCTTTAGAGAGAGACAGT 358

QY 144 TCCTTAGCAATTCCTACACTATGATGAAAGATATGCTCAAAATCCAGGGTTTTGTGGCT 203  
Db 359 TTTGCTACAATGCTGATATATATATATAAATAATGCTTTCCATGTGGATGGATTGATGGCT 418

QY 204 CATTGGTGGAAATTTGTTTGAATACAGAGAGGCAATTCAGCTAAGGCTGGAAGATACA 263  
Db 419 CTTTGGTGAATTCATGTTTAGAAGACCGAAGAACTACATCAACAGGTTGGAGATACA 478

QY 264 AAAACCTCGAAGAAACGCTTTCTCCAGAAAATTTACCAACATCAATTTGCCATCTT 323  
Db 479 AGAGCCCTGATGAGCACCCTTTCTCCGAGGATCTGCTCAACACCTGTTGCCACCTC 538

QY 324 ACTCCTTCAACAGTTTTTTCATCTCGTCTGCTTCAATTAACATAACAGTCCATCT 383  
Db 539 TCAGTATCCAAAGGTTTTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598

QY 384 GGAATATGATTTTCAAGAGTTACTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 443  
Db 599 GGAATATGATTTTGTATGAGCTTCTTCCAAAGATTAGTGAAGAGAGGATGATGTAAT 658

QY 444 ATCGGCAAACTGACGATTAATGACCTTTCAATTTGAGGTCATCTCTAGAGGATTCAT 503  
Db 659 ATGGATCCAGTGCTCTTTGTGACGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718

QY 504 ATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGCTCCTCAAGACTACAGGCTTTAA 563  
Db 719 ATGGTAAATTTTGGCAGAGGCTTAAATTCAGAGTCTCTGAGCTTACATGCTCGGA 778

QY 564 TTCGAGCTAAGGATAAGAGGATTGATGAATTTGTTGACATTTACAGCGTTGAAGAGA 623  
Db 779 TAATAGCAGGAGCTCGGATCAACTAATGCACTCTCTCACTATGAAACGGTGGAGCGTG 838

Search completed: September 25, 2004, 02:51:17  
Job time : 557 secs

QY 624 CGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGGTGTGGGCAAGTGAATCTTTAACA 683  
Db 839 CTAATTGAACATAGGGTGAAGCTAAGGCTAAGATCTTTGGACAGAGGTTAGGCG 898

QY 584 GTGATGACATGACAAAT---GAAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGT 740  
Db 899 CTGAAGACAAACGGCGCTCCACCAATGTACAAGATAAGGCCACGTTTGGTGGCTGAACCTGT 958

QY 741 ACAAAAATTGGGTGATACCTCTCCACCAAGAGGTTTCAGGTTGAGTACCTTCTCGCCGTC 800  
Db 959 ACAGCTACAGGATCATGCGCTAACCAAGAGGTTTCAAGTAGCTTACTTGTCTGAGGAGAT 1018

QY 801 TAGACTGAAGCATTACAATGCA 823  
Db 1019 TGGATTGATTGTTTACGATTGTA 1041

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2004, 23:12:46 ; Search time 3094 seconds

(without alignments)  
9844.675 Million cell updates/sec

Title: US-10-624-061-15

Perfect score: 1020

Sequence: 1 gcacgagagtttaattcaa.....atcttcaaaaaaaaaa 1020

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_estfun:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	541	53.0	560	10 BE821647	BE821647 GM700015A
2	525	51.5	563	10 AW666427	AW666427 sk36d12.y
3	511	50.1	512	12 EG383365	EG383365 sac19f08.
4	477	46.8	477	9 AI973566	AI973566 sc88e02.y

5	469	46.0	620	10 BF009709	BF009709 ss83b10.y
6	468.4	45.9	639	10 BE660542	BE660542 2-B12 Gna
7	457.6	44.9	648	10 AW774625	AW774625 EST333776
8	438.2	43.0	696	12 B1273174	B1273174 NF091B12f
9	431.2	42.3	654	10 AW775842	AW775842 EST334907
10	416.8	40.9	605	10 BF649665	BF649665 NF081H11E
11	415.6	40.7	422	12 BG239446	BG239446 sab70d03.
12	404	39.6	560	12 RM093460	RM093460 saj09c02.
13	400.4	39.3	529	14 CB827129	CB827129 LJNEST70f
14	388.8	38.1	605	12 B1420527	B1420527 LJNEST58b
15	388.8	38.1	726	14 CA800612	CA800612 saul9a05.
16	321	31.5	524	14 CB828627	CB828627 LJNEST88g
17	313.2	30.7	531	14 CB828445	CB828445 LJNEST86f
18	313.2	30.7	557	14 CB829043	CB829043 LJNEST93e
19	310.2	30.4	527	10 BE800482	BE800482 sg93h09.y
20	308.2	30.2	516	14 CB827444	CB827444 LJNEST74h
21	303.2	29.7	501	14 CB828062	CB828062 LJNEST82c
22	302.8	29.7	549	14 CB828614	CB828614 LJNEST88f
23	299.6	29.4	529	14 CB826768	CB826768 LJNEST66f
24	298	29.2	541	14 CB829238	CB829238 LJNEST95g
25	283.6	27.8	748	14 CD824727	CD824727 BN25.058J
26	271.6	26.6	466	10 AW720060	AW720060 LJNEST13b
27	262	25.7	459	12 B1419161	B1419161 LJNEST29b
28	258	25.3	885	13 BU693059	BU693059 SFD0085.P
29	251	24.6	698	10 BE577489	BE577489 L48-2335T
30	249.2	24.4	710	14 CF475712	CF475712 RTW2.11
31	249	24.4	394	10 AW774873	AW774873 EST334024
32	242.2	23.7	810	14 CD427426	CD427426 SAL.30.H0
33	241.4	23.7	427	12 B1420751	B1420751 LJNEST51c
34	239.4	23.5	1222	11 AY103806	AY103806 Zea mays
35	237.8	23.3	1196	11 AY107907	AY107907 Zea mays
36	236	23.1	660	13 BX253798	BX253798 BX253798
c 37	230.4	22.6	867	14 CB619882	CB619882 OSIIEA04L
38	229.4	22.5	680	13 CA080410	CA080410 SCVPAM105
39	227.8	22.3	821	14 CK196770	CK196770 FGAS00523
40	227.4	22.3	427	12 BG359738	BG359738 sac28a03.
c 41	216.6	21.2	725	14 CF486896	CF486896 POL1.40.B
42	213.4	20.9	696	14 CD919148	CD919148 G608.112E
c 43	212.4	20.8	607	9 AI486636	AI486636 EST244957
44	211	20.7	798	14 CB619881	CB619881 OSIIEA04L
45	210.8	20.7	623	14 CD998155	CD998155 QB88e10.x

## ALIGNMENTS

RESULT 1  
BE821647/c  
LOCUS BE821647 560 bp mRNA linear EST 24-MAY-2001  
DEFINITION GM700015A10A11 Gm-r1070 Glycine max cDNA clone Gm-r1070-5781 3', mRNA sequence.  
ACCESSION BE821647  
VERSION BE821647.1 GI:10253881  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
REFERENCE 1 (bases 1 to 560)  
AUTHORS Vodka, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corryell, V., Erpelting, J., Raph, C., Shoop, E., Pardinas, J., Liu, L. and Lewin, H.  
TITLE A Functional Genomics Program for Soybean (NSF 9872565)  
JOURNAL Unpublished (1999)  
COMMENT Other ESTs: AI973566 corresponding to Gm-cl019-435 (5')  
Contact: Vodka, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147



subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XLRP digested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Kelm and Dr. Virginia Coryell."

## ORIGIN

Query Match 51.5%; Score 525; DB 10; Length 563;  
Best Local Similarity 96.2%; Pred. No. 7.1e-108; Indels 0; Gaps 0;  
Matches 537; Conservative 0; Mismatches 21;

QY 94 GGTAGGCAAGAGATACCATCATTTATGGTCTCATTCAGAGAGCCAAAGTTCCTAGCAA 153  
Db 5 GGTAGGCAAGAGATACCATCATTTATGGTCTCATTCAGAGAGCCAAAGTTCCTAGCAA 64

QY 154 TTCTCACACCTATGATGAAGAATGCTCAATCCAGGTTTGTGGCTCATTTGGGA 213  
Db 65 TCCTCACACCTATGATGAAGAATGCTCAATCCAGGTTTGTGGCTCATTTGGTAGA 124

QY 214 ATTTGTTTGAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATACAAAAACCTGA 273  
Db 125 ATTTGTTTGAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATACAAAAACCTGA 184

QY 274 AGAAACGCCCTTCTCCAGAAAAATTTACCAACCATCAATTTGTGCCATCTTACTCTCAA 333  
Db 185 AGAAACGCCCTTCTCCAGAAAAATTTACCAACCATCAATTTGTGCCATCTTACTCTCAA 244

QY 334 ACAGTTTTCATCTGCTGCTCTCAATTAACATAAACAAGTCCATCTGGAAAAATGTA 393  
Db 245 ACAGTTTTCATCTGCTGCTCTCAATTAACATAAACAAGTCCATCTGGAAAAATGTA 304

QY 394 TTTCAAAGAGTTACTTCCATTTGCTTGTGCTTTCGGGTGATGAGCAACTATGCGCAAC 453  
Db 305 TTTCAAAGAGTTACTTCCATTTGCTTGTGCTTTCGGGTGATGAGCAACTATGCGCAAC 364

QY 454 TGCAGCTAATGACCTTTTATTTGCTGCTCATCTCTAGAGGATTCACATGAGAGTT 513  
Db 365 TGCAGCTAATGACCTTTTATTTGCTGCTCATCTCTAGAGGATTCACATGAGAGTT 424

QY 514 TGTAGCTGAGGTGAAATTCAGGGATGCTCTCAAGACTACGAGCTTTAATTCAGCTAA 573  
Db 425 TGTAGCTGAGGTGAAATTCAGGGATGCTCTCAAGACTACGAGCTTTAATTCAGCTAA 484

QY 574 GGATAAAGAGGATGATGAATTTGTTGACATTTTACAAAGCCTTGAAGAGCGGTGAGAA 633  
Db 485 GGATAAAGAGGATGATGAATTTGTTGACATTTTACAAAGCCTTGAAGAGCGGTGAGAA 544

QY 634 GAGAGTTCAAAGAGGCG 651  
Db 545 AGAGTTCAAAGAGGCG 562

## RESULT 3

BG363365 512 bp mRNA linear EST 28-NOV-2001  
LOCUS sac19f08.v1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE  
DEFINITION ID: Gm-cl051-2943 5' similar to TR:Q987H4 Q987H4 CHORISMATE MUTASE  
; mRNA sequence.

## ACCESSION

BG363365

## VERSION

BG363365.1 GI:13252454

## KEYWORDS

EST.

## SOURCE

Glycine max (soybean)

## ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

## REFERENCE

## AUTHORS

1. (bases 1 to 512)  
Shoemaker,R., Kelm,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.

## TITLE

## JOURNAL

## COMMENT

Public Soybean EST Project

Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: estewatson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 397.

## FEATURES

## source

1..512

Location/Qualifiers

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-2943"

/tissue\_type="floral meristematic mRNA"

/lab\_host="DH10B"

/clone\_lib="Gm-cl051"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from floral

meristematic mRNA provided by Dr. Halina Knap of Clemson

University. Complementary DNA was synthesized from mRNA

using a primer consisting of a poly(dT) sequence with a

XhoI restriction site. EcoRI adapters were ligated to the

blunt-ended cDNA fragments followed by XhoI digestion. The

cDNA fragments were directionally cloned into the

EcoRI-XhoI restriction site of the pBluescript vector. The

ligated cDNA fragments were transformed into DH10B host

cells (GibcoBRL). This library was constructed in the

laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 50.1%; Score 511; DB 12; Length 512;  
Best Local Similarity 99.8%; Pred. No. 1e-104; Indels 0; Gaps 0;  
Matches 511; Conservative 0; Mismatches 1;

QY 39 AACAAAGTCTGATTCCTGGGAATGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTA 98  
Db 1 AACAAAGTCTGATTCCTGGGAATGTACACGCTAGCTTCTGTGAGAGAGGATTTGGTTA 60

QY 99 GGCAAGAGGATACCATCAATTTATGGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAATTC 158  
Db 61 GGCAAGAGGATACCATCAATTTATGGTCTCATTTGAGAGAGCCAAAGTTCCTAGCAATTC 120

QY 159 ACACCTTATGATGAAGAATGATGCTCAAAATCCAGGTTTTTGTGGCTCATTTGGTGAATTTG 218  
Db 121 ACACCTTATGATGAAGAATGATGCTCAAAATCCAGGTTTTTGTGGCTCATTTGGTGAATTTG 180

QY 219 TTGTTAAGAAATACAGAGGCCATTCAGCTTAAGCTGGAAGATACAAAAACCTCGAAGAAA 278  
Db 181 TTGTTAAGAAATACAGAGGCCATTCAGCTTAAGCTGGAAGATACAAAAACCTCGAAGAAA 240

QY 279 ACACCTTCTTCCAGAAAATTTACCAACCATCAATTTGTCGCACTTACTCTCTCAACAGT 338  
Db 241 ACACCTTCTTCCAGAAAATTTACCAACCATCAATTTGTCGCACTTACTCTCTCAACAGT 300

QY 339 TTTTGCATCCTGCTGCTGCTTCAATTAACATAAACAAGTCCATCTCGAAAAATGATTTCA 398  
Db 301 TTTTGCATCCTGCTGCTGCTTCAATTAACATAAACAAGTCCATCTCGAAAAATGATTTCA 360



South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
High quality sequence stop: 404.

FEATURES

source  
1. .620  
Location/Qualifiers  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Williams"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl064-692"  
/tissue\_type="seedling epicotyls"  
/dev\_stage="2 week old"  
/lab\_host="DH10B"  
/clone\_lib="Gm-cl064"  
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
XhoI. The cDNA library was constructed from mRNA isolated  
from the epicotyls of 2 week old seedling in a growth  
chamber, excised above the soil level, and the plants  
were placed in a 100 ppm solution of auxin for 24 hours  
prior to harvesting. Complementary DNA was synthesized  
from mRNA using a primer consisting of a poly(dT)  
sequence with a XhoI restriction site. EcoRI adapters  
were ligated to the blunt-ended cDNA fragments followed  
by XhoI digestion. The cDNA fragments were directionally  
cloned into the EcoRI-XhoI restriction site of the  
pBluescript vector. The ligated cDNA fragments were  
transformed into DH10B host cells (GibcoBRL). This  
library was constructed in the laboratory of Dr. Randy  
Shoemaker."

ORIGIN

Query Match 46.0%; Score 469; DB 10; Length 620;  
Best Local Similarity 91.7%; Pred. No. 2.9e-95;  
Matches 518; Conservative 0; Mismatches 45; Indels 2; Gaps 2;  
Qy 8 AGCTTAATTTCAAGATGGCCAAAGCAGACGAGCAAAAGTCTCATTTCTGGGAATGTGAC 67  
Db 57 ACCTTAATTTCAAGATGGCCAAAGCAGACGAGCAAAAGTCTCATTTCTGGGAATGTGAC 116  
Qy 68 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTC 127  
Db 117 ACGTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGCTCTC 176  
Qy 128 ATTGAGAGAGCAAGTTCCTTAGCAATTCACACCTATGATGAAAGATGCTCAATC 187  
Db 177 ATTGAGAGAGCAAGTTCCTTAGCAATTCACACCTATGATGAAAGATGCTCAATC 236  
Qy 188 CAGGTTTTTGTGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAAGCT 247  
Db 237 CAGGTTTTTGTGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAAGCT 296  
Qy 248 AAGCTGGAAGATACAAAAACCCCTGAAGAAAAAGCCCTTCTCCAGAAAAATTTACACCA 307  
Db 297 AAGCTGGAAGATACAAAAACCCCTGAAGAAAAAGCCCTTCTCCAGAAAAATTTACACCA 356  
Qy 308 TCAATTTGCCATCTTACTCTCTCAACAGTTTTTGGTATCTGCTGCTCTCAATTAAC 367  
Db 357 TCAATTTGCCATCTTACTCTCTCAACAGTTTTTGGTATCTGCTGCTCTCAATTAAC 416  
Qy 368 ATAAACAGTCCATCTGGAAGAATGATTTCAAGAGATTAATTCATTTGCTTACTTCG 427  
Db 417 ATAAACAGTCCATCTGGAAGAAT-TAATTTCAAGAGTACTTTCAATTTGCTTACTTCG 475  
Qy 428 GGTATGATGCAACTATGCAACTGCACTATGCACTTATGCACTTATGCACTTATGCACT 487  
Db 476 GGGATGATGCAACTATGCACTTATGCACTTATGCACTTATGCACTTATGCACTTATGCACT 535  
Qy 488 TCTA-GAAGGATTCACATGGAAGTTTGTAGCTGAGGTGAAATTCAGGGATGCTCTCA 546  
Db 536 TTTAGGAGATTTCCCTTGGAAAGTTTGTACCTGAGGTGAAATTCAGGATTTCTCTCC 595  
Qy 547 AGACTAGAGCCTTTAATTCGAGCT 571

Db 596 AAACCTAGAGCCTTATTTCGAGCT 620  
RESULT 6  
BE660542  
LOCUS 2-B12 GmaxSC Glycine max cDNA, mRNA linear EST 06-SEP-2000  
DEFINITION BE660542  
ACCESSION BE660542  
VERSION BE660542.1 GI:9986434  
KEYWORDS EST  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.  
1 (bases 1 to 639)  
Harris, N., Chapman, B.P. and Gijzen, M.  
Gene expression in developing soybean seed coats  
Unpublished (2000)  
Contact: Gijzen M  
Agriculture and Agri-Food Canada  
1391 Sandford Street, London, Ontario, Canada N5V 4T3  
Tel: 519 457 1470  
Fax: 519 457 3997  
Email: gijzenm@agr.ca.  
Location/Qualifiers  
1. .639  
/organism="Glycine max"  
/mol\_type="mRNA"  
/cultivar="Harosoy 63"  
/db\_xref="taxon:3847"  
/tissue\_type="Seed coats"  
/lab\_host="E. coli strain XL0LR"  
/clone\_lib="GmaxSC"  
/note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI; This  
cDNA library was constructed from polyA+ enriched mRNA  
from green seed coats in mid to late developmental  
stage, average fresh weight 250 mg per seed. Traces of  
pod and embryo tissue also present. Complementary DNA was  
synthesized from mRNA using an XhoI-poly(dT)  
linker-primer. EcoRI adapters and the products were digested  
blunt-ended cDNA fragments and the products were ligated  
with XhoI for directional cloning into lambda ZAP Express  
vector. This lambda library was amplified once using E.  
coli host strain XL1 Blue MRF'. Inserts were then  
subcloned by mass excision using ExAssist helper phage for  
conversion into phagemid vector pBK-CMV in E. coli host  
strain XL0LR."  
Query Match 45.9%; Score 468.4; DB 10; Length 639;  
Best Local Similarity 90.5%; Pred. No. 4e-95;  
Matches 523; Conservative 0; Mismatches 51; Indels 4; Gaps 2;  
Qy 437 GGCACATATGCGCAAACTGCGAGCTAATGACCTTTTCATTTTGCAGTCCATCTAGAAGG 496  
Db 1 GGAACACTATGCAAACTGCGAGCTAATGATCTTTTCATTTTGCAGGCCATCTAGAAGG 60  
Qy 497 ATTCATATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAG 556  
Db 61 ATTCATATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCAAGACTACGAG 120  
Qy 557 CCTTTAATTCGAGCTAAGGATAAAGAGATTGATGAAATTTTGCATTTTCAAGCGTT 616  
Db 121 CCTTTAATTCGAGCTAAGGATAAAGAGATTGATGAAATTTTGCATTTTCAAGCGTT 180  
Qy 617 GAAGACGCTGAGGAGAGAGTTTGAAGAGAGGCTGCTGCTTTGGGAGAGAGTGAAT 676  
Db 181 GAAGACGCTGAGGAGAGAGTTTGAAGAGAGGCTGCTGCTTTGGGAGAGAGTGAAT 240  
Qy 677 CTTACAGCTGATGACAAATGACAAATGAAACCGTAAATTTGATCCTCAGTGGCTTTTACG 736

```

Db- 241 CTTGACAAATGAGCAGATGATGAGAAACCAATGTTTGTATCCATCAGTGGCTCTAGC 300
Qy 737 TTGTACAAAATTTGGGTGATACCTCTCACCAGGAGGTTTCCAGTTGAGTACCTCTTGGC 796
Db 301 TTGTACAAAATTTGGGTGATACCTCTTACCAGGAGGTTTCCAGTTGAGTACCTTGGC 360
Qy 797 CGTCTAGACTGAAGCATTACATGCAATGAGTTTGAAGAATGGAAGATGAATATGA 856
Db 361 CGTCTAGACTGAAGCATTATATGAGTTTGAAGAATTTTACAAGAATGGAAGTGAATATGA 420
Qy 857 TGTGTGTTTCAATGATTAAGCTCTTAAGTGTATCTTTATTTGCAACTTCAATGAGCTG 916
Db 421 TGTGTGTTTCAATGATTAAGCTCTTAAGTGTATCTTTATTTGCAACTTCAATGAGCTG 480
Qy 917 TTGATTCAGAAATATT-ATTGTAGCTATAGATTGTTACCTTTATTTTCAATGAGCTTTAT 975
Db 481 TTGATTTAAAAATATTAAATTTGAGCTTTAGATTGTTACTTTTATTTTCAATGAGCTTTAT 540
Qy 976 T---ACGAAAAGGCATTATTATATCATGATCTTCAAAA 1010
Db 541 TAGAAGAAAAAAGGATTAGATCAATGAACTTCAATA 578

```

RESULT 7  
 AW774625  
 LOCUS  
 DEFINITION  
 EST333776 KV3 Medicago truncatula cDNA clone pKV3-23M13, mRNA  
 sequence.  
 AW774625  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Medicago truncatula (barrel medic)  
 Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 648)  
 VandenBosch,K., Hurt,J., Moore,J., Beremand,P., Peng,H., Ellis,L.,  
 Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and  
 Fraser,C.M.  
 ESTs from roots of Medicago truncatula after Rhizobium inoculation  
 Unpublished (1999)  
 Contact: VandenBosch K  
 Department of Plant Biology  
 University of Minnesota  
 220 BioSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
 Tel: 612 624 2755  
 Fax: 612 625 1738  
 Email: kvandenb@ccs.umn.edu  
 Texas A&M EST name: T258044e  
 TIGR sequence name: MTESC79TK  
 More information is available at:  
 http://chryslie.tamu.edu/medicago  
 Seq primer: SKmod (CTA GAA CTA gtc gat CC).

Location/Qualifiers  
 1..648  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultivar="genotype A17"  
 /db\_xref="taxon:3880"  
 /clone="pKV3-23M13"  
 /tissue\_type="Seedling roots"  
 /dev\_stages="3 days post-inoculation with Sinorhizobium  
 meliloti"  
 /lab\_host="E. coli strain XLOLR"  
 /clone\_lib="KV3"

FEATURES  
 source  
 1..648  
 /note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
 XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA  
 was directionally ligated into the Unizap XR vector from  
 Stratagene and packaged using Gigapack III Gold packaging  
 extracts. Plasmids containing cDNA inserts were excised

from the recombinant lambda-Zap phage using Ex-assist  
 helper phage and propagated in XLOLR cells."

# ORIGIN

Query Match 44.9%; Score 457.6; DB 10; Length 648;  
 Best Local Similarity 81.6%; Pred. No. 1.1e-92; Indels 0; Gaps 0;  
 Matches 529; Conservative 0; Mismatches 119;

Qy 122 GGCTCATTTGAGAGAGCCAAAGTTCCCTAGCAATTTCTCACACCTATATGATGAAAAGTATGCT 181  
 Db 1 GGCTTATTGAGAGATCAAAAGTTTCCCTTTAATTTCTCAAACTATATGATCAGAAATATCTT 60  
 Qy 182 CAATCCAGGGTTTGTGGCTCATTTGGTGGAAATTTGTTGTTAAGAAATACAGAGGCCATT 241  
 Db 61 CAATCCCTGGTTTGTGGCTCATTTGGTGGAAATTTGTTGTTTACCAATATCTGAGATTGTT 120  
 Qy 242 CAAGCTAAGGCTCGAAGATACAAAACCCCTGAAGAAAACGCCCTTCTTCCAGAAAATTTA 301  
 Db 121 CAAGCTAAGGCTCGAAGATACAAAACCCCTGAAGAAAATCCCTTCTTCCAGAAAATTTA 180  
 Qy 302 CCACCATCAATTTGTGCATCTTACTCTTCAACACAGTTTTCGATCCTGGTGTGCTTCA 361  
 Db 181 TCATTGTCACTTGTTCATCTTACCCCTTCAAAAAGTTTTCGATCCTGGAGCGGCTCG 240  
 Qy 362 ATTAACATAAACAAAGTCCATCTGGAATAATGTTTCAAAAGAGTTTACTTCCATTGCTGCT 421  
 Db 241 ATTAACATAAACAAAGTCCATCTGGAATAATGTTTCAAAAGAGTTTACTTCCATTGCTGCT 300  
 Qy 422 ACTTCGGGTGATGATGGCAACTATGCGCAAACTGCGAGCTTAATGACCTTTTCAATTTGCG 481  
 Db 301 GCTTCCGGTGTGATGGCAACTATGCGCAAACTGCGAGCTAGCGACCTCTCAATTTGTCAG 360  
 Qy 482 TCCATCTCTAGAGAGGATTCACCTATGGAAGTTTGTAGCTGAGGTGAATTCAGGATGCT 541  
 Db 361 GCATTTCTAAGAGAGTTTCAATTTGGAAGTTTGTAGCTGAAAGTGAATTCAGGAAATCT 420  
 Qy 542 CCTCAAGACTACGAGCCTTTAATTTGAGCTAAGGATAAAGAGGATTGATGAAATTTGTTG 601  
 Db 421 CCTCAAGACTATGAGCCTTTAATTTGAGCTAAGGATAAAGAGGTTGATGAAATTTGTTG 480  
 Qy 602 ACATTTACAGGCTTGAAGAGCGGTGAGAGAGAGTTTGAAGAAGAGCTGTGTGTTT 661  
 Db 481 ACATTTAAGAGTGTGAAGAGATGTTAAGAAGAGGTTGAAGAAGAGGCAACATTTGTTT 540  
 Qy 662 GGGCAGGAAGTGAATCTTAACAGTGTATGACATGACATGAAACCGTAAATTTTATCCCA 721  
 Db 541 GGGCAGGATATAAGCCTTAACAGCAGTGTATGACATGAAAGTAAACAAAGTTTATCCCA 600  
 Qy 722 TCAGTGGCTTCTAGCTTGTACAAAATTTGGGTGATACCTCTTACCACCAAG 769  
 Db 601 TCAGTGGCTTCTAAGTTGTATGAGAAATGGGTAAATACCTCTTACTTAAG 648

## RESULT 8

BI273174  
 LOCUS  
 DEFINITION  
 NF091B12FL1F1096 Developing flower Medicago truncatula cDNA clone  
 NF091B12FL 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS

BI273174.1 GI:14883109

EST.

Medicago truncatula (barrel medic)

ORGANISM

Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.

1 (bases 1 to 696)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,  
 Flores,H.R., Inman,J.T., Weiler,J.W. and May,G.D.  
 Expressed sequence tags from the Samuel Roberts Noble Foundation  
 Medicago truncatula flower library  
 Unpublished (2001)



## COMMENT

Contact: May GD  
Plant Biology Division  
The Samuel Roberts Noble Foundation  
2510 Sam Noble Parkway, Ardmore, OK 73402, USA  
Tel: 580 221 7391  
Fax: 580 221 7380  
Email: gdmay@noble.org  
Insert Length: 696 Std Error: 0.00  
Plate: 091 row: B column: 12  
Seq primer: TCACACAGGAACAGCTATGAC.

## FEATURES

source  
1..696  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/db\_xref="taxon:3880"  
/clone="NF091B12FL"  
/tissue\_type="Developing flowers"  
/dev\_stage="Developmentally pooled. Contains a mixture of very young, developing, fully-opened flowers and flowers in early transition into pods."  
/clone\_lib="Developing flower"  
/note="Vector: Lambda Zap; cDNA was prepared from polyA+ enriched, pooled samples of equivalent amounts of total RNA from very young, developing, fully-opened flowers and flowers transitioning into pods. The cDNA was directionally ligated into the Uni-Zap XR vector (Stratagene) and packaged using the Gigapack III Gold packaging extracts. Phagemids containing cDNA inserts were in vivo excised from the recombinant Uni-Zap XR vector using ExAssist helper phage and the E. coli strain XL1-Blue MRF' (Stratagene). Excised plasmids were plated using SOLR cells."

## ORIGIN

Query Match 43.0%; Score 438.2; DB 12; Length 696;  
Best Local Similarity 79.0%; Pred. No. 2.5e-88;  
Matches 543; Conservative 0; Mismatches 142; Indels 2; Gaps 2;  
Qy 173 AAGTATGCTCAATCCAGGGTTTGTGGCTCAATGCTGAATTTGTTTGAAGATACA 232  
Db |||||  
Qy 2 AATTATCTTCAATCCCTGTTTTGTGGCTCAATGCTGAATTTGTTTCAACCAATAC 61  
Db |||||  
Qy 233 GAGGCCATTCAGTGAAGCTGGAAGATACAAAACCTCGAAGAAACGGCTTCTCCCA 292  
Db |||||  
Qy 62 GAGATTGTTCAAGTAAGCTGGAGATACAAAACCTCGAAGAAATCCCTCTTCCCA 121  
Db |||||  
Qy 293 GAAATTTACCAACATCAATGTGCCATCTACTCTTCAACAGTTTTCGATCTGGT 352  
Db |||||  
Qy 122 GAAATTTATCATGTGTCTGTTCCATCTTACCCCTTCAAAAGTTTTCGATCTCGA 181  
Db |||||  
Qy 353 GCTGCTTCAATACATAACAGTCCATCTGGAATGTATTTCAAGAGTTACTTCCA 412  
Db |||||  
Qy 182 GCCGCTTCGATTAACATAACAGTCTTATGGAATATGACTTTTATGACCTGCTTCCA 241  
Db |||||  
Qy 413 TTGCTTGCTACTTCGGGTGATGATGGCAACTATCGCAAACTGCAGCTAATGACCTTCA 472  
Db |||||  
Qy 242 TTGTTGCTGCTTCGGGTGATGATGGCAACTATGCAAACTGCAGCTAGCACTTCA 301  
Db |||||  
Qy 473 TTATGAGTCCATCTCTAGAGGATTCATATGGAAGTTTGTAGCTGAGGTGAATTC 532  
Db |||||  
Qy 302 TTGTTGAGGCGCAATTTCTAAAGAGTTTCAATATGGAAGTTTGTAGCTGAGGTGAATTC 361  
Db |||||  
Qy 533 AGGATGCTCTCAAGACTACAGCCCTTAATTCGAGCTAAGGATAAGAGGATTTGATG 592  
Db |||||  
Qy 362 AGGAAATCTCTCAAGACTATGACCTTTAATTCGTTCTAAGCAAGAGGGTTGATG 421  
Db |||||  
Qy 593 AAATGTTGACATTTACAAGCGTTTGAAGAGCGTGAAGAGAGAGTTTGAAGAAAGGCT 652  
Db |||||  
Qy 422 AAATGTTGACATTTAAGAGTGTGAAGAGATGTTAAAGAAAGGGTTGAAGAAAGGCA 481  
Db |||||  
Qy 653 GTGCTGTTGGCAGGAGTGAATCTTAACAGTATGACATGACATGAACACCGCTAA 712  
Db |||||  
Qy 482 ACATGTTTGGCAGGAGTAAAGCCTTAACAGCAGTGTGACAGTAAAGGTAAACAAAG 541  
Db |||||

## Qy

713 TTGTATCCATCAGTGGCTTCTAGCTTTGACAAAATTTGGTGATACCTCTCACCAGGAG 772

## Db

542 TTGTATCCATCAGTGGCTTCTAGCTTTGACAAAATTTGGTGATACCTCTCACCAGGAG 601

## Qy

773 GTTC-AGGTGAGTACCT-CTTGCCCGCTCTAGACTGAAGGCAATTCACATGAGTTGAA 830

## Db

602 GGTCAAGGTGAGTACCTTACTGCCCCCGTTGGATTGAACCACTATTATTTCATGTA 661

## Qy

831 TTTAGAAGATCGAAGATGATATGAT 857

## Db

662 TTTCAGAAAGACATATAATATCAT 688

## RESULT 9

AW775842

## LOCUS

EST334907 DSIL Medicago truncatula cDNA clone pDSIL-3E8, mRNA

## DEFINITION

sequence.

## ACCESSION

AW775842

## VERSION

AW775842.1 GI:7765655

## KEYWORDS

EST.

## SOURCE

Medicago truncatula (barrel medic)

## ORGANISM

Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

## REFERENCE

1 (bases 1 to 654)

## AUTHORS

Fedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gantt, G.S., Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Fraser, C.M.

## TITLE

ESTs from leaves of Medicago truncatula after inoculation with Colletotrichum trifolii

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Deborah A. Samac

Department of Plant Pathology

University of Minnesota

495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA

Tel: 612 625 1243

Fax: 651 649 5058

Email: debbys@umn.edu

Minnesota sequence name: M259175e

TIGR sequence name: MTEAJ28TK

More information is available at:

http://chrysis.tamu.edu/medicago

Seq primer: SKmod (CTA GAA CTA Gtg gat CC).

## FEATURES

Location/Qualifiers

1..654

/organism="Medicago truncatula"

/mol\_type="mRNA"

/cultivar="genotype A17"

/db\_xref="taxon:3880"

/clone="pDSIL-3E8"

/tissue\_type="leaves infected with Colletotrichum trifolii"

/dev\_stage="cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii"

/lab\_host="E. coli strain XL0LR"

/clone\_lib="DSIL"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-Zap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XL0LR cells. Note: EST may be of fungal origin."

## ORIGIN

Query Match 42.3%; Score 431.2; DB 10; Length 654;



```

ORGANISM      Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 422)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 402.
Location/Qualifiers
1..422
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl032-2453"
/tissue_type="unknown"
/lab_host="unknown"
/clone_lib="Unknown Library Type"
/notes="Vector: unknown; Site 1: unknown; Site 2: unknown;
This library is assigned to clones that are of uncertain
origin due to an error during re-array, prepping, and/or
analysis. Samples assigned this library are still genuine
Glycine Max cDNA sequence, but we are unsure of its source
and identification. Due to the clone id being
incorrectly assigned, THESE CLONES ARE NOT AVAILABLE TO
ORDER."

FEATURES             source
Query Match          40.7%; Score 415.6; DB 12; Length 422;
Best Local Similarity 99.1%; Pred. No. 3.3e-83;
Matches 418; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 357 CTTCAATTACATAAACAAGTCCATCTGGAAGATGATTTCGAAGAGTTACTTCCATTGC 416
|||
1 CTTCAATTACATAAACAAGTCCATCTGGAAGATGATTTCGAAGAGTTACTTCCATTGC 60
Db

QY 417 TTGCTACTTCGGGTGATGATGCGCAACTATGCGCAACTGAGCTAATGACCTTTTCATTAT 476
|||
61 TTGCTACTTCGGGTGATGATGCGCAACTATGCGCAACTGAGCTAATGACCTTTTCATTAT 120
Db

QY 477 TGCAGTCCATCTCAGAGAGATTACATATGGAAGTTTGTAGCTAGGTGAATTCAGG 536
|||
121 TGCAGTCCATCTCAGAGAGATTACATATGGAAGTTTGTAGCTAGGTGAATTCAGG 180
Db

QY 537 ATGCTCTCTCAAGACTACGAGCCCTTTAATTCGAGCTAAGGATAAGAGATTGATGAAT 596
|||
181 ATGCTCTCTCAAGACTACGAGCCCTTTAATTCGAGCTAAGGATAAGAGATTGATGAAT 240
Db

QY 597 TGTTCATTTTACAAGCGTTGAAGAGACGCTGAGGAAGAGAGTTGAAAGAGCGGTGG 656
|||
241 TGTTCATTTTACAAGCGTTGAAGAGACGCTGAGGAAGAGAGTTGAAAGAGCGGTGG 300
Db

QY 657 TGTTCGGCAGGAAGTGAATCTTTAAGAGTATGACATGACATGAATGAATCCGTAATTTG 716
|||
301 TGTTCGGCAGGAAGTGAATCTTTAAGAGTATGACATGACATGAATGAATCCGTAATTTG 360
Db

QY 717 ATCCATCAGTGGCTTCTAGCTTTGTACAAAATTTGGGTGATACCTCTCACCAGGAGTTC 776
|||
361 ATCCATCAGTGGCTTCTAGCTTTGTACAAAATTTGGGTGATACCTCTCACCAGGAGTTC 420
|||
777 AG 778
|||
421 AG 422

RESULT 12
BM0933460
LOCUS
DEFINITION
BM0933460 560 bp mRNA linear EST 30-NOV-2001
saj09c02.y1 Gm-cl065 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl065-9724 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE
; mRNA sequence.
ACCESSION
BM0933460
VERSION
BM0933460.1 GI:17022426
KEYWORDS
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 560)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 421.
Location/Qualifiers
1..560
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl065-9724"
/tissue_type="germinating shoots"
/lab_host="DH10B"
/clone_lib="Gm-cl065"
/notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from mRNA isolated
germinating shoots of the cultivar Williams. The seeds
were allowed to germinate for 24 hours prior to being
cold stressed for 2 days at 4C. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with a XhoI restriction site. EcoRI
adapters were ligated to the blunt-ended cDNA fragments
directionally cloned into the EcoRI-XhoI restriction site
of the pBluescript vector. The ligated cDNA fragments were
transformed into DH10B host cells (GibcoBRL). This library
was constructed in the laboratory of Dr. Randy
Shoemaker."

ORIGIN
Query Match          39.6%; Score 404; DB 12; Length 560;
Best Local Similarity 87.0%; Pred. No. 1.3e-80;
Matches 487; Conservative 0; Mismatches 0; Indels 73; Gaps 1;

QY 384 GGAAAATGTATTTCAAAGAGTTACTTCCATTGCTTGTACTTCGGGTGATGATGGCAACT 443
|||

```

|||||  
1 GGAATGATTTTCAAGAGTTACTTCCATTTGCTGCTACTTCGGTGATGATGCAACT 60  
444 ATGCGCAAACTGACCTAATGACCTTTCATTA----- 475  
61 ATGCGCAAACTGACCTAATGACCTTTCATTAATTCAGGACCAACTCAAGATTTTGTAG 120  
476 -----TTGCACTCCATCTCT 490  
121 TCTCTAAACTAGTATTGATCTCTGACTGATCATTTCTATAAAACATTTGCACTCCATCTCT 180  
491 AGAAGGATTCACTATGGAAGTTTCTAGCTGAGTGGAATTCAGGATGCTCTCAAGAC 550  
181 AGAAGGATTCACTATGGAAGTTTCTAGCTGAGTGGAATTCAGGATGCTCTCAAGAC 240  
551 TACGAGCTTTAAATTCGAGTAAAGTAAAGAGGATTTGAATTTGTCACATTTACA 610  
241 TACGAGCTTTAAATTCGAGTAAAGTAAAGAGGATTTGAATTTGTCACATTTACA 300  
611 AGCGTTGAAGACACGCTGAGGAAGAGAGTTGAAAAAGAGGCTGTGTGTTTGGGACGAA 670  
301 AGCGTTGAAGACACGCTGAGGAAGAGAGTTGAAAAAGAGGCTGTGTGTTTGGGACGAA 360  
671 GTGAATCTTAAAGATGATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCT 730  
361 GTGAATCTTAAAGATGATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCT 420  
731 TCTAGCTTCTACAAAATTTGGGTGATCTCTCAACAGAGGTTTCAGGTTGAGTACCTC 790  
421 TCTAGCTTCTACAAAATTTGGGTGATCTCTCAACAGAGGTTTCAGGTTGAGTACCTC 480  
791 TTGCGCGCTGATGACGATGAGGATTAATCAATGAGTTAGATTTAGAAAGATGGAAGATGA 850  
481 TTGCGCGCTGATGACGATGAGGATTAATCAATGAGTTAGATTTAGAAAGATGGAAGATGA 540  
851 ATATGATGTTGTTGTTCAAA 870  
541 ATATGATGTTGTTGTTCAAA 560

RESULT 13  
CB827129  
LOCUS  
DEFINITION  
LJN5870f7r Lotus japonicus nodule library 5 and 7 week-old Lotus  
corniculatus var. japonicus cDNA 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Lotus corniculatus var. japonicus (Lotus japonicus)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;  
Lotus.  
1 (bases 1 to 529)  
Colebatch, G., Freund, S., Trevisan, B. and Udvardi, M.  
Lotus japonicus root nodule ESTs: tools for functional genomics  
Unpublished (2000)  
Contact: Udvardi MK  
Molecular Plant Nutrition  
Max Planck Institute of Molecular Plant Physiology  
Am Muehlenberg 1, 14476 Golm, Germany  
Fax: 49 331 567 8250  
Email: udvardi@pimp-golm.mpg.de  
Seq primer: T7  
High quality sequence stop: 529.  
Location/Qualifiers  
1. 529  
/organism="Lotus corniculatus var. japonicus"  
/mol\_type="mRNA"  
/cultivar="Gifu (B-129)"  
/db\_xref="taxon:34305"  
/dev\_stage="5 and 7 week-old plants"

/clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"  
/note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI. The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."  
Query Match 39.3%; Score 400.4; DB 14; Length 529;  
Best Local Similarity 85.3%; Pred. No. 8.5e-80;  
Matches 446; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
QY 183 AAATCAGGGTTTGTGGCTCATTTGGTGGAAATTTGTTTAAAGATACAGAGGCAATTC 242  
DB 6 AAATCCGGGTTTGTGGCTCATTTGGTGGAAATTTGTTTCAAAACACTAGAGGCAATTC 65  
QY 243 AAGCTAAGGCTGAAGATACAAAACCCCTGAAGAAACCCCTTCTCCAGAAAATTTTAC 302  
DB 66 AAGCTATGGCTGAAGATACAAAACCCCTGAAGAAATTTGCTTCTCCAGAAATTTAC 125  
QY 303 CACCATCAATTTGCCATCTTACTCTCTCAACAGATTTTGGTGGTGGTGGTGGTGGT 362  
DB 126 CATCACCAATTTGGCTCATTTACCCCTTCAACAGATTTTGGCTGGAGCTGGTCCA 185  
QY 363 TTAACATTAACAGATCCATCTGGAATGATTTTCAAGAGTACTTCCATTCCTGCTTGA 422  
DB 186 TTAACATTAACAGATCGATTTGGAAATTTCTTTGATGATGTTTCCATTTGGTTG 245  
QY 423 CTTTGGGCTGATGATGGCAACTATGCGCAAACTGCACTTAATGACCTTCTTATTTGCA 482  
DB 246 CGTCCGGTGTGATGCAACTATGCAAACTGCACTAGCATCTTTCATTTATTCAGG 305  
QY 483 CCACTCTAGAGGATCTACTATGGAAGTTTGTAGCTGAGTGAATTCAGGATGCTC 542  
DB 306 CAATCTCTAGAGGATTCATTTATGGAAGTTTGTAGCTGAGGCAATTCAGGATCTC 365  
QY 543 CTCAGACTACGAGCTTTAAATTCGAGCTAAGATTAAGAGGATTTGATGAAATTTTGA 602  
DB 366 CTCAGACTATGAGCTTTAAATTCGCTTGAAGACACAGAGGTTGATGAAATTTGA 425  
QY 603 CATTTCAAGCGTTGAAGAGACGCTGAGGAGAGATTTGAAAGAGGCTGTGGTGGT 662  
DB 426 CATTTCAAGAGTGTTCAGAGATGCTGATTAAGAGAGTTGAAAGAGGCTGGTGGT 485  
QY 663 GCGAGGAAGTGAATCTTAAACAGTGAATGACATGCAATGAAA 705  
DB 486 NGCAGGAGGTAGCCCTAAACAGTGAATGAAAGAGGATGAAA 528  
RESULT 14  
BI420527 605 bp mRNA linear EST 15-AUG-2001  
LOCUS  
DEFINITION  
LJN58758b9r Lotus japonicus nodule library 5 and 7 week-old Lotus  
corniculatus var. japonicus cDNA 5', mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Lotus corniculatus var. japonicus (Lotus japonicus)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae;  
Lotus.  
1 (bases 1 to 605)  
Colebatch, G., Freund, S., Trevisan, B. and Udvardi, M.  
Lotus japonicus root nodule ESTs: tools for functional genomics  
Unpublished (2000)  
Contact: Udvardi MK  
Molecular Plant Nutrition  
Max Planck Institute of Molecular Plant Physiology  
Am Muehlenberg 1, 14476 Golm, Germany  
Fax: 49 331 567 8250

Email: udvardi@pimp-golm.mpg.de

Seq primer: T7

High quality sequence stop: 605.

Location/Qualifiers

## FEATURES

source

1. 605  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="mRNA"  
 /cultivar="Gifu (B-129)"  
 /db\_xref="taxon:34305"  
 /dev\_stage="5 and 7 week-old plants"  
 /clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"  
 /note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

## ORIGIN

Query Match 38.1%; Score 388.8; DB 12; Length 605;  
 Best Local Similarity 84.2%; Pred. No. 3.4e-77;  
 Matches 436; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 48 CTGATTCGGGATGTGTACACCTAGCTTCTGTGAGAGAGGATTTGGTTAGGCAAGG 107  
 Db 86 CAGATCGAATGATGTGTTTCCCTGGAGTCTGTGAGGAGGATCTGGTTAAGCAAGAAG 145  
 QY 108 ATACCATCAATTATGTCTCATTTGAGAGAGCCAAAGTTCCTGACCAATTTCTCACACCTATG 167  
 Db 146 ATACCATGTTTTTGGTCTCATTTGAGAGGCTAGGTTCCCAAGATTATCACACCTATG 205  
 QY 168 ATGAAAAGTAGTCAAAATCCAGGGTTTGTGGCTCATTTGGTGGAAATTTGTTTGAAGA 227  
 Db 206 ATGAAACTACTGGAAATCCCGGGTTTTTGTGGCTCATTTGGTGGAAATTTGTTTCAA 265  
 QY 228 ATACAGAGGCAATCAAGCTTAAGCTGGAAGATACAAAACCTGGAAGAAAGCCCTTCT 287  
 Db 266 ACATGAGGCATTCAGCTATGCTTGGAAAGTACCAAAACCTGGAAGAAATGCTTCT 325  
 QY 288 TCCGAGAATTTACCAACATCAATTTGTGCTTCTTCTTCAACAGTTTGTGCATC 347  
 Db 326 TCCGAGAATTTACCAATCAATTTGTGCTTCTTCTTCAACAGTTTGTGCATC 385  
 QY 348 CTGCTGCTGCTCAATTAACATAAACAGTCCATCTGGAATGATTTCAAGAGTTAC 407  
 Db 386 CTGAGCTGCTTCCATTAACATAAACAGTCCATCTGGAATGATTTCAAGAGTTAC 445  
 QY 408 TTCCATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 467  
 Db 446 TTCCATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 505  
 QY 468 TTTCATTTGAGTCCATCTCTAGAGGATTCACATGGAAGTTTGTAGCTGAGGTGA 527  
 Db 506 TTTCATTTGAGGCAATCTCTAGAGGATTCACATGGAAGTTTGTAGCTGAGGTGA 565  
 QY 528 AATTGAGGATGCTCTCTCAAGACTAGGAGCTTTTAATTCG 567  
 Db 566 AATTGAGGATGCTCTCTCAAGACTAGGAGCTTTTAATTCG 605

## RESULT 15

CA800612

LOCUS

DEFINITION

CA800612 726 bp mRNA linear EST 05-DEC-2002  
 sau19a05.y1 Gm-cl062 Glycine max cDNA clone SOYBEAN CLONE ID:  
 Gm-cl062-8481 5' similar to TR:Q957H4 Q957H4 CHORISMATE MUTASE ;  
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

## REFERENCE

## AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Corryell,V.,  
 Khanna,A., Bolle,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project  
 Unpublished (1999)

## TITLE

## JOURNAL

## COMMENT

Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact: ccu@resgen.com web site:

www.resgen.com

Seq primer: -40RP from Gibco

High quality sequence stop: 409.

Location/Qualifiers

1. 726

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl062-8481"

/tissue\_type="stem tissue of greenhouse grown plants"

/dev\_stage="1 month old"

/lab\_host="DH10B"

/clone\_lib="Gm-cl062"

/note="Vector: Bluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from stem tissue of 1 month old greenhouse grown plants  
 for the cultivar Raiden. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 pCly(dr) sequence with a XhoI restriction site. EcoRI  
 adapters were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA fragments were  
 directionally cloned into the EcoRI-XhoI restriction site  
 of the pBluescript vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (GibcoBRL). This library  
 was constructed in the laboratory of Dr. Randy Shoemaker."

## ORIGIN

Query Match 38.1%; Score 388.8; DB 14; Length 726;  
 Best Local Similarity 98.2%; Pred. No. 3.4e-77;  
 Matches 393; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 241 TCAAGCTAAGGCTGGAAGATACAAAAACCCCTGAGAAAAGCCCTTCTTCCAGAAAAATTT 300  
 Db 327 TGATGCACTTGTCTGGAAGATACAAAAACCCCTGAGAAAAGCCCTTCTTCCAGAAAAATTT 386  
 QY 301 ACCACCATCAATTTGTCATCTTACTCTCTTCAAAACAGTTTTTGGATCCTCTGGTGTCTTC 360  
 Db 387 ACCACCATCAATTTGTCATCTTACTCTCTTCAAAACAGTTTTTGGATCCTCTGGTGTCTTC 446  
 QY 361 AATTAAACATAAACAAGTCCATCTGAAAAATGATTTCAAAGAGTTACTTCCATTTGCTTGC 420  
 Db 447 AATTAAACATAAACAAGTCCATCTGAAAAATGATTTCAAAGAGTTACTTCCATTTGCTTGC 506  
 QY 421 TACTTCGGGTGATGATGGCAACTATGCGAAACTGCAAGCTAATGACCTTTTCAATTATTCGA 480  
 Db 507 TACTTCGGGTGATGATGGCAACTATGCGAAACTGCAAGCTAATGACCTTTTCAATTATTCGA 566  
 QY 481 GTCCATCTCTAGAGGATTTCACATATGAAAAGTTTGTAGCTGAGGTGAATTCAGGGATGC 540  
 Db 567 GTCCATCTCTAGAGGATTTCACATATGAAAAGTTTGTAGCTGAGGTGAATTCAGGGATGC 626  
 QY 541 TCCTCAAGACTACGAGCCCTTTAATTCAGCTAAGGATAAAGAGGATTGATGAATTTGTT 600

us-10-624-061-15.rst

Mon Sep 27 07:44:33 2004

Db	627	TCCTCAAGACTACGAGCCCTTTAATTCGAGCTAAGATTAAGAGGATTTGATGAATTGTT	686
Qy	601	GACATTTACAAGCGTTGAAGAGACGCGTGAGGAAGAGATT	640
Db	687	GACATTTACAAGCGTTGAAGAGACGCGGAGGAAGAGATT	726

Search completed: September 25, 2004, 01:24:49  
Job time : 3099 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 01:24:55 ; Search time 74 Seconds  
(without alignments)  
996.553 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339  
Sequence: 1 MAXAAEQSPDSGNVYTLASV.....NWVPLTKVQVEYLLRLD 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339	100.0	261	6	ABU08094 Soybean c
2	818	61.1	265	5	AAM49654 A. thalia
3	818	61.1	265	5	ABB33424 Herbicida
4	818	61.1	265	6	ABU08098
5	818	61.1	330	5	AAM49655 Construct
6	698	52.1	334	5	ABB32474 Herbicida
7	696	52.0	334	6	ABU08097
8	691.5	50.9	257	6	ABU08093 Rice chor
9	673	50.3	132	6	ABU08089
10	670.5	50.1	331	6	ABU08092
11	669.5	50.0	336	5	ABB31535
12	602	45.0	207	6	ABU08095
13	599.5	44.8	319	3	AAG06003 Arabidops
14	599.5	44.8	325	3	AAG06002
15	599.5	44.8	347	3	AAG06001 Arabidops
16	489	36.5	154	6	ABU08090
17	424.5	31.7	263	6	AAO19566 Protein e
18	423	31.6	280	3	AAO19566 M sterili
19	147	11.0	93	6	ABU08087
20	131	9.8	62	6	ABU08098
21	105	7.8	360	6	ABU00704
22	103	7.7	781	4	AAU35676 Haemophil
23	103	7.7	781	6	ABU30650 Protein e
24	96.5	7.2	655	4	ABG25473 Novel hum
25	96	7.2	459	4	AAG2076 S. epider

26	96	7.2	459	4	AAG2730 S. epider
27	96	7.2	464	5	ABP39856 Staphyloc
28	95.5	7.1	594	6	ABM73376 Staphyloc
29	94	7.0	1037	4	ABG04999 Novel hum
30	93.5	7.0	739	6	ABU25591 Protein e
31	92.5	6.9	537	4	ABBS7939 Prosophil
32	92	6.9	605	6	ABBS7939 Protein s
33	91.5	6.8	433	6	ABR52685 Protein s
34	91.5	6.8	649	6	ADA20823 Rice stre
35	91.5	6.8	676	6	ABU25397 Protein e
36	90.5	6.8	554	2	AAW69223 HAP4 prot
37	90	6.7	693	6	ABU44329 Protein e
38	89.5	6.7	590	7	ADC95055 E. faeciu
39	88.5	6.6	680	6	ADB23150 Environme
40	88.5	6.6	856	3	AA341660 Human ORF
41	88.5	6.6	1173	4	AA97911 Human G-p
42	88.5	6.6	1948	5	ABP69390 Human pol
43	88.5	6.6	1958	7	ADE14371 Human int
44	88	6.6	1072	2	AAV29460 African h
45	87.5	6.5	596	6	ABU24103 Protein e

## ALIGNMENTS

### RESULT 1

ABU08094	ID	ABU08094	standard; protein; 261 AA.
XX	AC	ABU08094;	
XX	DT	22-MAY-2003 (first entry)	
XX	DE	Soybean chorismate mutase protein.	
XX	KW	Soybean; plant; enzyme; chorismate mutase; aromatic amino acid; diet;	
XX	KW	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;	
XX	KW	phenylalanine; prephenate dehydrogenase; herbicide; food crop.	
XX	OS	Glycine max.	
XX	PN	US2002184658-A1.	
XX	PD	05-DEC-2002.	
XX	PF	03-DEC-1999; 99US-00454279.	
XX	PR	04-DEC-1998; 98US-0110845P.	
XX	PA	(FALC/) FALCO S. C.	
XX	PA	(FAMO/) FAMODU O. O.	
XX	PI	(LEEJ/) LEE J.	
XX	PI	Falco SC, Famodu OO, Lee J;	
XX	DR	WPI: 2003-328651/31.	
XX	DR	N-PSDB; ABX93058.	

New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and nucleic acids, useful for facilitating design and/or identifying inhibitors of those enzymes that may be used as herbicides and for producing antibodies.

Claim 11; Fig 1; 32pp; English.

The invention discloses isolated polynucleotides encoding chorismate mutase polypeptides. Aromatic amino acids must be included in the diets of animals. In the aromatic amino acid biosynthetic pathway chorismate is converted to anthranilate during tryptophan biosynthesis and is converted to prephenate, the branch point for tyrosine and phenylalanine biosynthesis. Chorismate mutase catalyzes the conversion of chorismate to prephenate. Also disclosed are methods for selecting an isolated polynucleotide that affects the level of expression of a tyrosine

CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
CC evaluating compounds for their ability to inhibit the activity of a  
CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
CC targets for herbicides that will not affect animals, and overexpression  
CC of these enzymes may be used to increase the content of aromatic amino  
CC acid in food crops. The polypeptides may also be used to design and/or  
CC identify inhibitors of those enzymes that may be used as herbicides. The  
CC nucleic acids may be used to create transgenic plants, as probes for the  
CC genetic and physical mapping of the genes and as markers for traits  
CC linked to those genes. The sequence presented is a soybean chorismate  
CC mutase protein  
XX  
XX  
SQ Sequence 261 AA;

Query Match 100.0%; Score 1339; DB 6; Length 261;  
Best Local Similarity 100.0%; Pred. No. 2.2e-133;  
Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAKAAEQSPDSGNVYTLASVREDLVROEDTIYGLIERAKPSPNSHTYDEKYAQIQGFCG 60  
DB 1 MAKAAEQSPDSGNVYTLASVREDLVROEDTIYGLIERAKPSPNSHTYDEKYAQIQGFCG 60  
QY 61 SLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPSPSIVPSYFQFLHPGGAASININKSI 120  
DB 61 SLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPSPSIVPSYFQFLHPGGAASININKSI 120  
QY 121 WKMYFKELLPLATSGDDGNYAQTAANDLSLQISRIHYGKFAEVKFRDAPQDYEP 180  
DB 121 WKMYFKELLPLATSGDDGNYAQTAANDLSLQISRIHYGKFAEVKFRDAPQDYEP 180  
QY 181 IRAKDKGELMKLLTFTSVETVRKRVKKAUVFGQEVNLSDDNDNENRKFDPSPVASSLY 240  
DB 181 IRAKDKGELMKLLTFTSVETVRKRVKKAUVFGQEVNLSDDNDNENRKFDPSPVASSLY 240  
QY 241 KKWVIPLTKVEQVVEYLLRLD 261  
DB 241 KKWVIPLTKVEQVVEYLLRLD 261

RESULT 2  
ID AAM49654  
XX AAM49654 standard; protein; 265 AA.  
AC AAM49654;  
DT 17-MAY-2002 (first entry)  
XX A. thaliana cytosolic chorismate mutase-2 protein.  
DE Shikimate pathway; chorismate mutase; prephenate dehydrogenase;  
XX vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;  
KW feed; pharmaceutical; cosmetic; aromatic amino acid; salicylic acid;  
KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;  
KW tocopherol; tocotrienol.  
XX Arabidopsis thaliana.  
XX Arabidopsis thaliana.  
XX WO200200901-A1.  
XX 03-JAN-2002.  
XX 28-JUN-2001; 2001WO-EP007391.  
XX 29-JUN-2000; 2000DE-01030647.  
XX 21-DEC-2000; 2000DE-01064454.  
XX (SUNG-) SUNGENE GMBH & CO KGAA.  
XX Badur R, Geiger M, Kunze I, Sommer S;  
XX WPI; 2002-164442/21.  
XX

DR N-PSDB; ABA99613.  
XX Preparing fine chemicals, particularly Vitamins E and K, useful as  
PT antioxidants e.g. in foods or medicine, by growing organisms with altered  
PT shikimate biosynthesis pathway.  
XX  
PS Claim 9; Page 60-61; 74pp; German.  
XX This invention describes a novel method for the preparation of fine  
CC chemicals by culturing organisms in which the shikimate pathway has been  
CC altered relative to the wild type. The method involves a construct  
CC containing a plasmid transgene, a chorismate mutase and/or  
CC prephenate dehydrogenase, linked to at least one regulatory sequence for  
CC transcription or translation in plants. The method is used to produce  
CC chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
CC fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
CC also aromatic amino acids, salicylic or folic acid derivatives,  
CC phenylpropanoids, flavanoids etc., especially in transgenic plants. Also  
CC plants with increased Vitamin E contents have improved resistance to  
CC abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
CC shikimate pathway are useful as foods, fodder and in preparation of  
CC processed foodstuffs. Transgenic plants with modified shikimate pathways  
CC have increased content of desired chemicals, particularly tocopherols  
CC and/or tocotrienols. This sequence represents the Arabidopsis thaliana  
CC cytosolic chorismate mutase-2 protein used in the method of the invention  
XX  
XX  
SQ Sequence 265 AA;

Query Match 61.1%; Score 818; DB 5; Length 265;  
Best Local Similarity 61.1%; Pred. No. 5.3e-78;  
Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
QY 1 MAKAAEQSPDSG--NVYTLASVREDLVROEDTIYGLIERAKPSPNSHTYDEKYAQIQG 58  
DB 1 MARVFSDSGSGCNVLSLRLRESLROEDTIVFSLIERAKPPLNSPAPFESRCLDSGS 60  
QY 59 CGSLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPSPSIVPSYFQFLHPGGAASININK 118  
DB 61 FSSLTFFVRETEIIQAKGVREYEPENPFLENIPHSVFTHKYPSALHPKALSININK 120  
QY 119 SIWKMYFKELLPLATSGDDGNYAQTAANDLSLQISRIHYGKFAEVKFRDAPQDYEP 178  
DB 121 QIWDIVFKELLPLFVKPGDDGNYPSTAASDLACLQALSRIHYGKFAEVKFRDAPQDYEP 180  
QY 179 PLIRAKDKGELMKLLTFTSVETVRKRVKKAUVFGQEVNLSDDNDNENRKF--DPSVA 236  
DB 181 PAIRAQDREALMKLLTFEKEVENMKRVQKKAETFGQEVKNSGYGDESKKKYKVDPLLA 240  
QY 237 SSLLYKQWVIPLTKVEQVVEYLLRLD 261  
DB 241 SRVYGEWLIPLTKVEQVVEYLLRLD 265

RESULT 3  
ID ABB93424  
XX ABB93424 standard; protein; 265 AA.  
AC ABB93424;  
XX 31-MAY-2002 (first entry)  
XX Herbicidally active polypeptide SEQ ID NO 2635.  
DE Herbicidal; plant; agriculture; herbicide.  
XX Arabidopsis thaliana.  
XX Arabidopsis thaliana.  
XX WO200210210-A2.  
XX 07-FEB-2002.  
XX 28-AUG-2001; 2001WO-EP009892.  
XX







Db	255	MDMLTFTPTVEDAIKRVEMKTRTYGOEVKVGMEKEEBEENESHVYKISPIVLGDLYG	314
QY	242	NWVPLTKVQVEYLLRLD	261
Db	315	DWIMPLTKVQVEYLLRLD	334
RESULT 7			
ABU08097	ID ABU08097 standard; protein; 334 AA.		
XX	AC	ABU08097;	
XX	DT	22-MAY-2003 (first entry)	
XX	DE	A. thaliana chorismate mutase protein #1.	
KW	KW	Plant; enzyme; chorismate mutase; aromatic amino acid; diet;	
KW	KW	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;	
XX	XX	phenylalanine; prephenate dehydrogenase; herbicide; food crop.	
OS	OS	Arabidopsis thaliana.	
XX	PN	US2002184658-A1.	
XX	PD	05-DEC-2002.	
XX	PF	03-DEC-1999; 99US-00454279.	
XX	PR	04-DEC-1998; 98US-0110845P.	
XX	PA	(FALC/) FALCO S C.	
PA	PA	(FAMO/) FAMODU O O.	
XX	PA	(LEEJ/) LEE J.	
XX	PI	Falco SC, Famodu OO, Lee J;	
XX	DR	WPI; 2003-328651/31.	
XX	PT	New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and	
PT	PT	nucleic acids, useful for facilitating design and/or identifying	
PT	PT	inhibitors of those enzymes that may be used as herbicides and for	
XX	XX	producing antibodies.	
PS	PS	Example 3; Fig 1; 32pp; English.	
XX	CC	The invention discloses isolated polynucleotides encoding chorismate	
CC	CC	mutase polypeptides. Aromatic amino acids must be included in the diets	
CC	CC	of animals. In the aromatic amino acid biosynthetic pathway chorismate is	
CC	CC	converted to anthranilate during tryptophan biosynthesis and is converted	
CC	CC	to prephenate, the branch point for tyrosine and phenylalanine	
CC	CC	biosynthesis. Chorismate mutase catalyzes the conversion of chorismate to	
CC	CC	prephenate. Also disclosed are methods for selecting an isolated	
CC	CC	polynucleotide that affects the level of expression of a tyrosine	
CC	CC	biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid	
CC	CC	fragment encoding a tyrosine biosynthetic enzyme polypeptide and	
CC	CC	evaluating compounds for their ability to inhibit the activity of a	
CC	CC	tyrosine biosynthetic enzyme. The polypeptides can be used produce	
CC	CC	antibodies. Chorismate mutase and prephenate dehydrogenase are good	
CC	CC	targets for herbicides that will not affect animals, and overexpression	
CC	CC	of these enzymes may be used to increase the content of aromatic amino	
CC	CC	acid in food crops. The polypeptides may also be used to design and/or	
CC	CC	identify inhibitors of those enzymes that may be used as herbicides. The	
CC	CC	nucleic acids may be used to create transgenic plants, as probes for the	
CC	CC	genetic and physical mapping of the genes and as markers for traits	
CC	CC	linked to those genes. The sequence presented is the A. thaliana	
CC	CC	chorismate mutase protein #1	
XX	XX	Sequence 334 AA;	
XX	XX	Query Match	52.0%; Score 696; DB 6; Length 334;
XX	XX	Best Local Similarity	50.4%; Pred. No. 6.9e-65;

Matches	131;	Conservative	46;	Mismatches	75;	Indels	8;	Gaps	1;
QY	10	DSGNVYTLASVREDLVROEDTIIYGLIERAKFPNSHTYDEKYAQIOGFCGSLVEFVYKN	69						
Db	75	DESESLTLEGIRNSLIHQEDSIIIFGLIERAKYCYNADTYDPTAFDMGFGNSLVEYVWVG	134						
QY	70	TEAIQAKAGRYKNPENAAPPENIPPSIVPSYSFKQFLHPCGAASININKSIWKMYFKELL	129						
Db	135	TEKLHAKVGRFKSPDEHPFFDDLPPEPMLPQPKVLHFAADSININKKIWNMYFRDLV	194						
QY	130	PLLATSGDDGNYAOTAAANDLSLQISIRRIHYGKFAEVEKFRDAPQDYVEPLIRAKDKEGL	189						
Db	195	PRLVKGGDDGNYGSTAVCDALICQLSKRIHYGKFAEAKFQASPEAYESAIIKAQDKDRL	254						
QY	190	MKLITFTSVETVRKRVKKAIVFGQEVNLSDDNDNENR-----KFPDPSVASSLYK	241						
Db	255	MDMLTFTPTVEDAIKRVEMKTRTYGOEVKVGMEKEEBEENESHVYKISPIVLGDDLYG	314						
QY	242	NWVPLTKVQVEYLLRLD	261						
Db	315	DWIMPLTKVQVEYLLRLD	334						
RESULT 8									
ID	ABU08093	ID ABU08093 standard; protein; 257 AA.							
XX	AC	ABU08093;							
XX	DT	22-MAY-2003 (first entry)							
XX	DE	Rice chorismate mutase protein.							
XX	KW	Rice; plant; enzyme; chorismate mutase; aromatic amino acid; diet;							
KW	KW	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;							
XX	XX	phenylalanine; prephenate dehydrogenase; herbicide; food crop.							
OS	OS	Oryza sativa.							
XX	PN	US2002184658-A1.							
XX	PD	05-DEC-2002.							
XX	PF	03-DEC-1999; 99US-00454279.							
XX	PR	04-DEC-1998; 98US-0110845P.							
XX	PA	(FALC/) FALCO S C.							
PA	PA	(FAMO/) FAMODU O O.							
XX	PA	(LEEJ/) LEE J.							
XX	PI	Falco SC, Famodu OO, Lee J;							
XX	DR	WPI; 2003-328651/31.							
XX	DR	N-PSDB; ABX93057.							
XX	PPT	New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and							
PPT	PPT	nucleic acids, useful for facilitating design and/or identifying							
PPT	PPT	inhibitors of those enzymes that may be used as herbicides and for							
XX	XX	producing antibodies.							
PS	PS	Claim 11; Fig 1; 32pp; English.							
XX	CC	The invention discloses isolated polynucleotides encoding chorismate							
CC	CC	mutase polypeptides. Aromatic amino acids must be included in the diets							
CC	CC	of animals. In the aromatic amino acid biosynthetic pathway chorismate is							
CC	CC	converted to anthranilate during tryptophan biosynthesis and is converted							
CC	CC	to prephenate, the branch point for tyrosine and phenylalanine							
CC	CC	biosynthesis. Chorismate mutase catalyzes the conversion of chorismate to							
CC	CC	prephenate. Also disclosed are methods for selecting an isolated							
CC	CC	polynucleotide that affects the level of expression of a tyrosine							
CC	CC	biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid							
CC	CC	fragment encoding a tyrosine biosynthetic enzyme polypeptide and							

CC evaluating compounds for their ability to inhibit the activity of a  
 CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
 CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
 CC targets for herbicides that will not affect animals, and overexpression  
 CC of these enzymes may be used to increase the content of aromatic amino  
 CC acid in food crops. The polypeptides may also be used to design and/or  
 CC identify inhibitors of those enzymes that may be used as herbicides. The  
 CC nucleic acids may be used to create transgenic plants, as probes for the  
 CC genetic and physical mapping of the genes and as markers for traits  
 CC linked to those genes. The sequence presented is a rice chorismate mutase  
 CC protein  
 XX  
 SQ Sequence 257 AA;

Query Match 50.9%; Score 681.5; DB 6; Length 257;  
 Best Local Similarity 52.1%; Pred. NO. 1.6e-63;  
 Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;  
 QY 6 EQSPSGNVYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKYAQIOGFCGSLVEF 65  
 DB 1 KORIDQSEILTDNINTSLVRQEDSIIFSLERAQFCYNADIYDKNAFVDFGDSLVEF 60  
 QY 66 VVKNTAEIAQKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKMYF 125  
 DB 61 MYRETEKLHQGVRYKSPDEHPFFEDLPEPLLPQYPKVLHPADSNINKSIWKMYF 120  
 QY 126 KELLPLLATSGDDGNYAQTAAANDLSLQISIRHYGFVAEVRFPADQYELIRAKD 185  
 DB 121 DELLPLRVKESGDGNYGSSALCDTICLQALSRIHYGFVAEVRFPADQYELIRAKD 180  
 QY 186 KEGLMKLLFTVEETVRKVEKKAIVFGQEVNLSDDNENR-KFPDPSVASSLYKNWV 244  
 DB 181 CQQLMELLTYETVERAIEHREKAKIFGQEVNLSDDNENR-KFPDPSVASSLYKNWV 240  
 QY 245 IPLTKEVQVEYLLRLID 261  
 DB 241 MPLTKEVQVAYLLRLID 257

RESULT 9  
 ABU08089  
 ID ABU08089 standard; protein; 132 AA.  
 AC ABU08089;  
 DT 22-MAY-2003 (first entry)  
 XX Protein encoded by soybean chorismate mutase expressed sequence tag.  
 DE Soybean; plant; enzyme; chorismate mutase; expressed sequence tag; EST;  
 XX aromatic amino acid; diet; biosynthesis; chorismate; anthranilate;  
 KW tryptophan; prephenate; tyrosine; phenylalanine;  
 KW prephenate dehydrogenase; herbicide; food crop.  
 XX Glycine max.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 130  
 FT /note= "Encoded by GGN"  
 FT  
 FT  
 PN US2002184658-A1.  
 XX  
 XX 05-DEC-2002.  
 XX  
 XX 03-DEC-1999; 99US-00454279.  
 XX  
 XX 04-DEC-1998; 98US-0110845P.  
 XX  
 XX (FALC/) FALCO S C.  
 PA (FAMO/) FAMODU O O.  
 PA (LEEJ/) LEE J.  
 XX  
 XX Falco SC, Famodu OO, Lee J;  
 PI

XX WPI; 2003-328651/31.  
 DR N-PSDB; ABX93053.  
 DR  
 XX New tyrosine biosynthetic enzyme, chorismate mutase, proteins and  
 PT nucleic acids, useful for facilitating design and/or identifying  
 PT inhibitors of those enzymes that may be used as herbicides and for  
 PT producing antibodies.  
 XX  
 XX Claim 11; Page 16-17; 32pp; English.  
 PS  
 XX The invention discloses isolated polynucleotides encoding chorismate  
 CC mutase polypeptides. Aromatic amino acids must be included in the diets  
 CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
 CC converted to anthranilate during tryptophan biosynthesis and is converted  
 CC to prephenate, the branch point for tyrosine and phenylalanine  
 CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
 CC prephenate. Also disclosed are methods for selecting an isolated  
 CC polynucleotide that affects the level of expression of a tyrosine  
 CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
 CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
 CC evaluating compounds for their ability to inhibit the activity of a  
 CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
 CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
 CC targets for herbicides that will not affect animals, and overexpression  
 CC of these enzymes may be used to increase the content of aromatic amino  
 CC acid in food crops. The polypeptides may also be used to design and/or  
 CC identify inhibitors of those enzymes that may be used as herbicides. The  
 CC nucleic acids may be used to create transgenic plants, as probes for the  
 CC genetic and physical mapping of the genes and as markers for traits  
 CC linked to those genes. The sequence presented is the protein encoded by a  
 CC soybean chorismate mutase expressed sequence tag (EST)  
 XX  
 SQ Sequence 132 AA;

Query Match 50.3%; Score 673; DB 6; Length 132;  
 Best Local Similarity 98.5%; Pred. NO. 4.6e-63;  
 Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 10 DSGNVYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKYAQIOGFCGSLVEFVKN 69  
 DB 1 DSGNVYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKYAQIOGFCGSLVEFVKN 60  
 QY 70 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKMYFKELL 129  
 DB 61 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKMYFKELL 120  
 QY 130 PLLATSGDDGNY 141  
 DB 121 PLLATSGDDGNY 132

RESULT 10  
 ABU08092  
 ID ABU08092 standard; protein; 331 AA.  
 AC ABU08092;  
 XX  
 XX 22-MAY-2003 (first entry)  
 DT  
 XX Corn chorismate mutase protein.  
 DE  
 XX Corn; plant; enzyme; chorismate mutase; aromatic amino acid; diet;  
 KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
 KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.  
 XX  
 XX Zea mays.  
 OS  
 XX US2002184658-A1.  
 PN  
 XX 05-DEC-2002.  
 XX  
 XX 03-DEC-1999; 99US-00454279.  
 XX  
 XX Falco SC, Famodu OO, Lee J;  
 PI

```
XX PR 04-DEC-1998; 98US-0110845P.
XX PA (FALC/) FALCO S C.
XX PA (FAMO/) FAMODU O O.
XX PA (LEEU/) LEE J.
XX PI Falco SC, Famodu OO, Lee J;
XX WPI; 2003-328651/31.
XX DR N-PSDB; ABX93056.
XX PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and
XX PT nucleic acids, useful for facilitating design and/or identifying
XX PT inhibitors of those enzymes that may be used as herbicides and for
XX PT producing antibodies.
XX PS Claim 11; Fig 1; 32pp; English.
XX CC The invention discloses isolated polynucleotides encoding chorismate
XX CC mutase polypeptides. Aromatic amino acids must be included in the diets
XX CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is
XX CC converted to anthranilate during tryptophan biosynthesis and is converted
XX CC to prephenate, the branch point for tyrosine and phenylalanine
XX CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to
XX CC prephenate. Also disclosed are methods for selecting an isolated
XX CC polynucleotide that affects the level of expression of a tyrosine
XX CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid
XX CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and
XX CC evaluating compounds for their ability to inhibit the activity of a
XX CC tyrosine biosynthetic enzyme. The polypeptides can be used produce
XX CC antibodies. Chorismate mutase and prephenate dehydrogenase are good
XX CC targets for herbicides that will not affect animals, and overexpression
XX CC of these enzymes may be used to increase the content of aromatic amino
XX CC acid in food crops. The polypeptides may also be used to design and/or
XX CC identify inhibitors of those enzymes that may be used as herbicides. The
XX CC nucleic acids may be used to create transgenic plants, as probes for the
XX CC genetic and physical mapping of the genes and as markers for traits
XX CC linked to those genes. The sequence presented is a corn chorismate mutase
XX CC protein
XX SQ Sequence 331 AA;
Query Match 50.1%; Score 670.5; DB 6; Length 331;
Best Local Similarity 51.4%; Pred. No. 3.5e-62;
Matches 130; Conservative 48; Mismatches 74; Indels 1; Gaps 1;
QY 10 DSGNVYTLASVREDLVQRQEDTIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFPVK 69
DB 79 DRSDILTDLSIRQLKLEDSIIIFGLIERAQFCYNADTYDSNAFMDGFGGSLVEIVRE 138
QY 70 TEATQAKAGRYKNPEENAFFPENTPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 129
DB 139 TEKLHAQVGRYKSPDEHPFSSKLPPEPLPPMQYPRVLHPIDADSININKSIWKYFDELL 198
QY 130 PLLATSGDDGNYAQTANDLSLQSISSRIHYGKFAEVKFRDAPQDYEPILIRAKDKEGL 189
DB 199 PRLVKEGSDGNAGSALCDTCTCQALSRRIHYGKFAEVKFRDAPQDYEPILIRAKDQDL 258
QY 190 MKLITFTSVETVAKRVEKKAIVFGQEVNLSNDNEN-RKFDPSVASSLYKNWVIPLT 248
DB 259 MLLITYETVERAIEHRVEAKKIFGQEVNIGAKDNGSPVPYKIRPSLVAELYSYRIMPLT 318
QY 249 KEVOVEYLLRLD 261
DB 319 KEVEVAYLLKRLD 331
RESULT 11
ABB91515
ID ABB91515 standard; protein; 316 AA.
XX AC ABB91515;
```

```
XX 31-MAY-2002 (first entry)
XX DE Herbicidally active polypeptide SEQ ID NO 726.
XX KW Herbicidal; plant; agriculture; herbicide.
XX OS Arabidopsis thaliana.
XX FN WO200210210-A2.
XX PD 07-FEB-2002.
XX PF 28-AUG-2001; 2001WO-EP009892.
XX PR 28-AUG-2001; 2001WO-EP009892.
XX PA (FARB ) BAYER AG.
XX PI Tietjen K, Weidler M;
XX DR WPI; 2002-269010/31.
XX PT Identifying plant target proteins for herbicidally active compounds,
XX PT comprising aligning and comparing nucleic acid or amino acid sequences
XX PT from plant with nucleic acid or amino acid sequences from non-plant
XX PT organisms.
XX PS Claim 5; SEQ ID NO 726; 261pp + Sequence Listing; English.
XX CC The invention relates to identifying target proteins (ABB90790-ABB94016)
XX CC for herbicidally active compounds, comprising aligning and comparing
XX CC nucleic acid or amino acid sequences from plant with nucleic acid or
XX CC amino acid sequences from non-plant organisms using suitable search
XX CC parameters, where plant sequences having an E-value greater by a factor
XX CC of 3 than the E-value of most similar non-plant sequences are selected.
XX CC The polypeptides or nucleic acids encoding them are useful for
XX CC identifying modulators. The identified modulators are useful as
XX CC herbicides
XX SQ Sequence 316 AA;
Query Match 50.0%; Score 669.5; DB 5; Length 316;
Best Local Similarity 51.4%; Pred. No. 4.1e-62;
Matches 130; Conservative 44; Mismatches 78; Indels 1; Gaps 1;
QY 10 DSGNVYTLASVREDLVQRQEDTIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFPVK 69
DB 64 DESEYLVKLESIRHSLIROEDSIIFNLERAQRYNADTYDEDAFTMEGFCGSLVEFVRE 123
QY 70 TEATQAKAGRYKNPEENAFFPENTPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 129
DB 124 TEKLHAQVGRYKSPDEHPFSSKLPPEPLPPMQYPRVLHPIDADSININKSIWKYFDELL 183
QY 130 PLLATSGDDGNYAQTANDLSLQSISSRIHYGKFAEVKFRDAPQDYEPILIRAKDKEGL 189
DB 184 PRLVKEGSDGNAGSALCDTCTCQALSRRIHYGKFAEVKFRDAPQDYEPILIRAKDQDL 243
QY 190 MKLITFTSVETVAKRVEKKAIVFGQEVNLSNDNEN-RKFDPSVASSLYKNWVIPLT 248
DB 244 MOLLITYETVERAIEHRVEAKKIFGQEVNIGAKDNGSPVPYKIRPSLVAELYSYRIMPLT 303
QY 249 KEVOVEYLLRLD 261
DB 304 KEVOVEYLLRLD 316
RESULT 12
ABU08095
ID ABU08095 standard; protein; 207 AA.
XX AC ABU08095;
XX AC ABU08095;
```

DT	22-MAY-2003	(first entry)	
XX	DE	Wheat chorismate mutase protein.	
XX	XX	Wheat; plant; enzyme; chorismate mutase; aromatic amino acid; diet;	
KW	KW	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;	
KW	KW	phenylalanine; prephenate dehydrogenase; herbicide; food crop.	
XX	OS	Triticum aestivum.	
XX	PN	US2002184658-A1.	
XX	PD	05-DEC-2002.	
XX	PF	03-DEC-1999; 99US-00454279.	
XX	PR	04-DEC-1998; 98US-0110845P.	
XX	PA	(FALC/) FALCO S C.	
XX	PA	(FAMO/) FAMODU O O.	
XX	PA	(LEEJ/) LEE J.	
XX	PI	Falco SC, Famodu OO, Lee J;	
XX	DR	WPI; 2003-328651/31.	
XX	DR	N-PSDB; ABX93059.	
XX	PT	New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and	
XX	PT	nucleic acids, useful for facilitating design and/or identifying	
XX	PT	inhibitors of those enzymes that may be used as herbicides and for	
XX	XX	producing antibodies.	
PS	Claim 11; Fig 1; 32pp; English.		
CC	The invention discloses isolated polynucleotides encoding chorismate		
CC	mutase polypeptides. Aromatic amino acids must be included in the diets		
CC	of animals. In the aromatic amino acid biosynthetic pathway chorismate is		
CC	converted to anthranilate during tryptophan biosynthesis and is converted		
CC	to prephenate, the branch point for tyrosine and phenylalanine		
CC	biosynthesis. Chorismate mutase catalyses the conversion of chorismate to		
CC	prephenate. Also disclosed are methods for selection of a tyrosine		
CC	polynucleotide that affects the level of expression of a tyrosine		
CC	biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid		
CC	fragment encoding a tyrosine biosynthetic enzyme polypeptide and		
CC	evaluating compounds for their ability to inhibit the activity of a		
CC	tyrosine biosynthetic enzyme. The polypeptides can be used produce		
CC	antibodies. Chorismate mutase and prephenate dehydrogenase are good		
CC	targets for herbicides that will not affect animals, and overexpression		
CC	of these enzymes may be used to increase the content of aromatic amino		
CC	acid in food crops. The polypeptides may also be used to design and/or		
CC	identify inhibitors of those enzymes that may be used as herbicides. The		
CC	nucleic acids may be used to create transgenic plants, as probes for the		
CC	genetic and physical mapping of the genes and as markers for traits		
CC	linked to those genes. The sequence presented is a wheat chorismate		
CC	mutase protein		
XX	Sequence 207 AA;		
XX	Query Match	45.0%; Score 602; DB 6; Length 207;	
XX	Best Local Similarity	57.8%; Fred. No. 3.2e-55;	
XX	Matches 118; Conservative	27; Mismatches 53; Indels 6; Gaps 2;	
QY	64	EFVVKNTAEIQAAGRYKXNPENAEFFPENLPSPISVPSFKQFLHPCGAASINIKSTWKM	123
Db	4	EFFVRAEVLHAKAGHYKQKEDVFFSQDLPPPPVPTKGRPKVLPFASLVCYNDAIWKM	63
QY	124	YFKLLPILATSGDGNVQAQTAANDLSLQISRIHYGKFAVSKFRDAPQDYELPIRA	183
Db	64	YFNELLPLFTADGGDGNVAFVALDFACLQALSRIHCGKYVAEVPKDPQDYSPIRA	123
QY	184	KDKEGLMKLLTFTSVEETVRKRVKKAQVVFQEVNLT---NSDDND--NENRKFDSVAS	237
Db	124	KDTNALMNLITFTAVEEKVKRVEKKARIFGQNTVLEDSVKGQDGDACDCHCKVDKVL	183
QY	238	SLYKNWVPLTKVEVQVEYLLRLD	261
Db	184	KLYDMVMVPLTKDVEVEYLLRLD	207
RESULT 13			
AAG06003			
ID	AAG06003	standard; protein; 319 AA.	
XX	AC	AAG06003;	
XX	XX	17-OCT-2000 (first entry)	
XX	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 2617.	
KW	KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	XX	termination sequence.	
OS	Arabidopsis thaliana.		
XX	PN	EP1033405-A2.	
XX	PD	06-SEP-2000.	
XX	XX	25-FEB-2000; 2000EP-00301439.	
XX	XX	25-FEB-1999; 99US-0121825P.	
PR	05-MAR-1999; 99US-0123180P.		
PR	03-MAR-1999; 99US-0123548P.		
PR	23-MAR-1999; 99US-0125788P.		
PR	25-MAR-1999; 99US-0126264P.		
PR	29-MAR-1999; 99US-0126785P.		
PR	01-APR-1999; 99US-0127462P.		
PR	06-APR-1999; 99US-0128234P.		
PR	08-APR-1999; 99US-0128714P.		
PR	16-APR-1999; 99US-0129845P.		
PR	19-APR-1999; 99US-0130077P.		
PR	21-APR-1999; 99US-0130449P.		
PR	23-APR-1999; 99US-0130510P.		
PR	23-APR-1999; 99US-0130891P.		
PR	28-APR-1999; 99US-0131449P.		
PR	30-APR-1999; 99US-0132048P.		
PR	30-APR-1999; 99US-0132407P.		
PR	04-MAY-1999; 99US-0132484P.		
PR	05-MAY-1999; 99US-0132485P.		
PR	06-MAY-1999; 99US-0132486P.		
PR	07-MAY-1999; 99US-0132487P.		
PR	11-MAY-1999; 99US-0132863P.		
PR	14-MAY-1999; 99US-0134218P.		
PR	14-MAY-1999; 99US-0134219P.		
PR	14-MAY-1999; 99US-0134221P.		
PR	14-MAY-1999; 99US-0134370P.		
PR	18-MAY-1999; 99US-0134768P.		
PR	19-MAY-1999; 99US-0134941P.		
PR	20-MAY-1999; 99US-0135124P.		
PR	21-MAY-1999; 99US-0135353P.		
PR	24-MAY-1999; 99US-0135629P.		
PR	25-MAY-1999; 99US-0136021P.		
PR	27-MAY-1999; 99US-0136392P.		
PR	28-MAY-1999; 99US-0136782P.		
PR	01-JUN-1999; 99US-0137222P.		
PR	03-JUN-1999; 99US-0137523P.		
PR	04-JUN-1999; 99US-0137502P.		
PR	07-JUN-1999; 99US-0137724P.		
PR	08-JUN-1999; 99US-0138094P.		
PR	10-JUN-1999; 99US-0138540P.		
PR	10-JUN-1999; 99US-0138847P.		
PR	14-JUN-1999; 99US-0139119P.		
PR	16-JUN-1999; 99US-0139452P.		
PR	16-JUN-1999; 99US-0139453P.		

PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	30-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	22-JUN-1999;	99US-0139899P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140353P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140695P.	PR	10-SEP-1999;	99US-0153070P.
PR	24-JUN-1999;	99US-0140823P.	PR	13-SEP-1999;	99US-0153758P.
PR	28-JUN-1999;	99US-0140991P.	PR	15-SEP-1999;	99US-0154018P.
PR	29-JUN-1999;	99US-0141287P.	PR	16-SEP-1999;	99US-0154039P.
PR	30-JUN-1999;	99US-0141842P.	PR	20-SEP-1999;	99US-0154779P.
PR	01-JUL-1999;	99US-0142154P.	PR	22-SEP-1999;	99US-0155113P.
PR	02-JUL-1999;	99US-0142055P.	PR	23-SEP-1999;	99US-0155486P.
PR	06-JUL-1999;	99US-0142390P.	PR	24-SEP-1999;	99US-0156596P.
PR	08-JUL-1999;	99US-0142803P.	PR	28-SEP-1999;	99US-0156458P.
PR	09-JUL-1999;	99US-0142920P.	PR	29-SEP-1999;	99US-0156596P.
PR	12-JUL-1999;	99US-0142877P.	PR	04-OCT-1999;	99US-0157117P.
PR	13-JUL-1999;	99US-0143542P.	PR	05-OCT-1999;	99US-0157533P.
PR	14-JUL-1999;	99US-0143624P.	PR	06-OCT-1999;	99US-0157865P.
PR	15-JUL-1999;	99US-0144005P.	PR	07-OCT-1999;	99US-0158032P.
PR	16-JUL-1999;	99US-0144085P.	PR	08-OCT-1999;	99US-0158232P.
PR	18-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144632P.	PR	18-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144841P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145132P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160989P.
PR	23-JUL-1999;	99US-0145224P.	PR	25-OCT-1999;	99US-0161404P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161403P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145918P.	PR	26-OCT-1999;	99US-0161359P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145931P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0145938P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			

Query Match 44.8%; Score 599.5; DB 3; Length 319;

Best Local Similarity 48.7%; Pred. No. 1.2e-54;

Matches 113; Conservative 45; Mismatches 71; Indels 3; Gaps 1;

QY 10 DSGNVTLASVREDLVROEDTIIYGLIERAKPFSNHYDEKIAQIOFCGSLVFVUKN 69

Db 75 DESESLTLEGIRNSLIROEDSIIFGLIERAKYCNADTYDPTAFDMDFGNSLVEYMKG 134

QY 70 TEAIQAKAGRYKNPENAFPPENLPPSIVPSYSFKQFLHPGAASININKSIWKYFKELL 129

Db 135 TELHAKVGRFKSPDEHPFDDLPBPMLPPLQYPKVLHFAADSIININKKIWNMYFRDIV 194  
QY 130 PLIATSGDDGNVAQTAANDLSLOSISRIHVGKFAVEKFRDAPQDYEPLIRAKDKGL 189  
Db 195 PRUVKGGDGNVGTAVCDALICQLSKRIHYGKFVAEAKFQASPEAYESAIAKQDKDAL 254  
QY 190 MKLLTFTSVEETVRKEVEKAVVFGQEVNLNSDDNDNENRKPDPSPVASSLYK 241  
Db 255 MDMLTPTVEDAIKKRVEKTRTYGQEVKVGMEKEEBEEBEGN---ESHVYK 303

RESULT 14  
AAG06002  
ID AAG06002 standard; protein; 325 AA.  
XX  
AC AAG06002;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2616.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 18-MAY-1999; 99US-0134370P.  
PR 19-MAY-1999; 99US-0134766P.  
PR 20-MAY-1999; 99US-0134941P.  
PR 21-MAY-1999; 99US-0135124P.  
PR 24-MAY-1999; 99US-0135353P.  
PR 25-MAY-1999; 99US-0135629P.  
PR 27-MAY-1999; 99US-0136021P.  
PR 28-MAY-1999; 99US-0136382P.  
PR 01-JUN-1999; 99US-0136782P.  
PR 03-JUN-1999; 99US-0137222P.  
PR 04-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137502P.  
PR 08-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 08-JUL-1999; 99US-0142390P.  
PR 09-JUL-1999; 99US-0142803P.  
PR 12-JUL-1999; 99US-0142920P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144083P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.



PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148585P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 07-SEP-1999; 99US-0151930P.  
PR 10-SEP-1999; 99US-0152363P.  
PR 13-SEP-1999; 99US-0153070P.  
PR 15-SEP-1999; 99US-0153758P.  
PR 16-SEP-1999; 99US-0154018P.  
PR 20-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 23-SEP-1999; 99US-0155139P.  
PR 24-SEP-1999; 99US-0155486P.  
PR 28-SEP-1999; 99US-0155659P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159339P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 21-OCT-1999; 99US-0159594P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160788P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161932P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.  
PR 29-OCT-1999; 99US-0162143P.

Query Match 44.8%; Score 599.5; DB 3; Length 325;  
Best Local Similarity 48.7%; Pred. No. 1.2e-54;  
Matches 113; Conservative 45; Mismatches 71; Indels 3; Gaps 1;  
10 DSGNVYTLASVRELDVQEETIYGLERAKFFSNHSHTYDEKYAQIQGFCGSLVEEVVKN 69

Db 81 DESESLTEGIRNSLRQEDSIIFGLERAKYCYNADTYDPTAFDMDGFGSLVEYVWG 140  
Qy 70 TEAIQAKAGYKXPEENAPENLPSPISVPSYFKQFLHPGAASININKSINKMYEKELL 129  
Db 141 TEKHAQVGRFKSPDEHPFFDDLPPEMLPPQYQPKVLHFAADSININKKINWNYFRDLV 200  
Qy 130 PLLATSGDDGNYAQTAAANDLSLIQSISSRIHYGKFAEVKFRDAPQOYEPLIRAKDKKGL 189  
Db 201 PRLVKGDDGNYGTVACDAICLCQCLSKRIHYGKFAEAKFQASPEAYESAIAQDDKDAL 260  
Qy 190 MKLITTSVEETVRKEVEKAVVFGGVNLNSDDNENKFKDPSPVASSLYK 241  
Db 261 MDMLTFTVEDAIKRVEMKTRTYGQEVKVGMEKEEEEEEGN---ESHVYK 309

RESULT 15  
AAG06001  
ID AAG06001 standard; protein; 347 AA.  
XX AAG06001;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2615.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 09-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 17-JUN-1999; 99US-0139453P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 23-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144008P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145132P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149829P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157533P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158332P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0158293P.  
PR 13-OCT-1999; 99US-0158294P.  
PR 13-OCT-1999; 99US-0158295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159337P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 44.8%; Score 599.5; DB 3; Length 347;  
Best Local Similarity 48.7%; Pred. No. 1.3e-54;  
Matches 113; Conservative 45; Mismatches 71; Indels 3; Gaps 1;

QY	10	DSGNVYTLASVREDLVROEDTITVGLIRAKFPNSHTYDEKVAQIQFGSLVEFVYKN	69
DB	103	DESESLTLEGIRNSLIRQEDSILFGLERAKYCVNADTYDPTAPDMDFNGSLVEYVYKG	162
QY	70	TEAIOAKAGRYNPPENAFPPENLPPIVPSYSFKQFLHPGASININKSIWKMYFKELL	129
DB	163	TEKHAAXGRFSPDEHPFPDDLPBPMLPPIQYPKVLHPAADSININKIIMNMYFRDLV	222
QY	130	PILATSGDDGNVAQTANDLSLLQISIRKIHGKFAVAFPRDAQDYEPILIRAKDEGL	189
DB	223	PLVYKKGDDGNYSTAVCDALICLOCLSKRIHYGKFVAEAKFQASPEAYESAIRKQDDAL	282
QY	190	MKLTFETSVETVRKRVKKAIVFGQEVNLNSDDNENNRKFPDPVASSLYK	241
DB	283	MDVLTFFPYVEDAIXGRVEMKTRTYQGEYKVGMEKEKEEEN--ESHVYK	331

Search completed: September 25, 2004, 02:52:38  
Job time : 77 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004. Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:51:21; Search time 23 seconds  
(without alignments)  
585.843 Million cell updates/sec

Title: US-10-624-061-16  
Perfect score: 1339  
Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NNVPLTKREVQVEYLLRLD 261

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pap:\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pap:\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pap:\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pap:\*  
5: /cgn2\_6/prodata/2/iaa/PCUTS COMB.pap:\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339	100.0	261	4	US-09-454-279-16
2	818	61.1	265	4	US-09-454-279-22
3	696	52.0	334	4	US-09-454-279-21
4	681.5	50.9	257	4	US-09-454-279-14
5	673	50.3	132	4	US-09-454-279-6
6	670.5	50.1	331	4	US-09-454-279-12
7	602	45.0	207	4	US-09-454-279-18
8	489	36.5	154	4	US-09-454-279-8
9	147	11.0	93	4	US-09-454-279-2
10	131	9.8	62	4	US-09-454-279-4
11	96	7.2	464	4	US-09-134-001C-4701
12	96	7.2	753	4	US-09-543-681A-5022
13	92	6.9	946	3	US-09-074-579-3
14	92	6.9	946	3	US-09-388-774-3
15	90.5	6.8	554	3	US-09-319-989-6
16	89.5	6.7	590	4	US-09-107-532A-4682
17	89	6.5	808	4	US-09-543-681A-4995
18	86.5	6.5	930	4	US-09-134-001C-5314
19	86.5	6.5	930	4	US-09-386-962C-10
20	83.5	6.2	1138	1	US-07-973-320-2
21	83	6.2	612	1	US-08-344-695-2
22	83	6.2	769	4	US-09-540-236-3110
23	83	6.2	902	4	US-09-107-532A-3837
24	83	6.2	1070	2	US-08-633-770A-2
25	83	6.2	1070	4	US-08-280-197-6
26	81	6.0	851	4	US-09-071-035-326
27	81	6.0	851	4	US-09-071-035-330

28	81	6.0	851	4	US-09-071-035-334	Sequence 334, Appl
29	81	6.0	946	4	US-09-546-153-1	Sequence 1, Appli
30	81	6.0	962	4	US-09-134-000C-4257	Sequence 4257, Ap
31	80.5	6.0	1180	3	US-09-224-024-28	Sequence 28, Appl
32	80.5	6.0	1180	5	PCT-US94-07902-28	Sequence 28, Appl
33	80.5	6.0	1440	3	US-09-357-251-37	Sequence 37, Appl
34	80.5	6.0	1512	4	US-09-443-184-48	Sequence 48, Appl
35	80	6.0	1110	1	US-08-118-441-29	Sequence 29, Appl
36	80	6.0	1110	3	US-08-338-579A-29	Sequence 29, Appl
37	80	6.0	1110	5	PCT-US94-09851-29	Sequence 3764, Ap
38	79.5	5.9	438	4	US-09-107-532A-3764	Sequence 2, Appli
39	79.5	5.9	535	4	US-08-969-046-2	Sequence 7, Appli
40	79.5	5.9	998	2	US-08-415-788-7	Sequence 3, Appli
41	79.5	5.9	1029	2	US-08-415-788-3	Sequence 12, Appl
42	79.5	5.9	1792	4	US-09-561-818A-12	Sequence 10, Appl
43	79.5	5.9	1816	4	US-09-561-818A-10	Sequence 1, Appli
44	79	5.9	997	1	US-08-232-540-1	Sequence 1, Appli
45	79	5.9	997	1	US-08-428-949A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-454-279-16  
; Sequence 16, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Lee, Jian-Ming  
; FILE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; CURRENT APPLICATION NUMBER: US/09/454, 279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-454-279-16

Query Match	100.0%	Score	1339;	DB	4;	Length	261;
Best Local Similarity	100.0%	Pred. No.	2e-137;				
Matches	261;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAKAAEQSPDSGNVYTLASVREDLVRQEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCG	60				
Db	1	MAKAAEQSPDSGNVYTLASVREDLVRQEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCG	60				
Qy	61	SLVEFVVKNTAIQAKAGRYKNPKNPEENAFPPENLPSIVPSYFKQFLHPGASININKSI	120				
Db	61	SLVEFVVKNTAIQAKAGRYKNPKNPEENAFPPENLPSIVPSYFKQFLHPGASININKSI	120				
Qy	121	WQVYFKEILLPLATSGDDGNVACQTAANDLSLQSTISRIHYGKFAEVKFRDAPQDYEP	180				
Db	121	WQVYFKEILLPLATSGDDGNVACQTAANDLSLQSTISRIHYGKFAEVKFRDAPQDYEP	180				
Qy	181	IRAKDKKGLMKLLTTSVEETVRKVEKKAUVFQGVNLSDDNDNENKRDPSVASSLY	240				
Db	181	IRAKDKKGLMKLLTTSVEETVRKVEKKAUVFQGVNLSDDNDNENKRDPSVASSLY	240				
Qy	241	KNWVPLTKREVQVEYLLRLD	261				
Db	241	KNWVPLTKREVQVEYLLRLD	261				

RESULT 2  
US-09-454-279-22  
; Sequence 22, Application US/09454279

Patent No. 6627798  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming  
 TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 FILE REFERENCE: BB1299 US NA  
 CURRENT APPLICATION NUMBER: US/09/454,279  
 CURRENT FILING DATE: 1999-12-03  
 EARLIER APPLICATION NUMBER: 60/110,845  
 EARLIER FILING DATE: 1998-12-04  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 22  
 LENGTH: 265  
 TYPE: PRT  
 ORGANISM: Arabidopsis thaliana  
 US-09-454-279-22

Query Match 61.1%; Score 818; DB 4; Length 265;  
 Best Local Similarity 61.1%; Pred. No. 1e-80;  
 Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
 QY 1 MAKAAEQSDSG--NVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIOGF 58  
 Db 1 MARVFESDGSNCNLSLDLRESLIRQEDTIVFSLIERAKFPLNSPAPFESRCLDSGS 60  
 QY 59 CGSLVEFVVVNTAIAQAKAGRYKNEENAFPPENLPPSIVPSYSGFKQFLHPCAASININK 118  
 Db 61 FSSLTEFFVRETEIIQAKGRVEYEENPFLENIPHSVFTTHKYPSSALHPKALSVINK 120  
 QY 119 STWKVFKELLPLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRDAPQDYE 178  
 Db 121 QIWDIYFKELLPLFKVKGDDGNYPSTAAASDLACLQALSRIHYGKFAEVRDAPQDYE 180  
 QY 179 PLIRAKDEGLMKLITFTSVETVRKVEKKAIVFGQVNLNSDDNENRKE--DPSVA 236  
 Db 181 PAIRAQDREALMKLITFEKVEWVKRQKAEIFGQVEKNSGYGDESKKKYKVDPLLA 240  
 QY 237 SLLYKNWVPLTKVEQVYLLRLD 261  
 Db 241 SRIYGEWLIPLTKLVEVYLLRLD 265

RESULT 3  
 US-09-454-279-21  
 Query Match 52.0%; Score 696; DB 4; Length 334;  
 Best Local Similarity 50.4%; Pred. No. 2.8e-67;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming  
 TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 FILE REFERENCE: BB1299 US NA  
 CURRENT APPLICATION NUMBER: US/09/454,279  
 CURRENT FILING DATE: 1999-12-03  
 EARLIER APPLICATION NUMBER: 60/110,845  
 EARLIER FILING DATE: 1998-12-04  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 21  
 LENGTH: 334  
 TYPE: PRT  
 ORGANISM: Arabidopsis thaliana  
 US-09-454-279-21

Query Match 52.0%; Score 696; DB 4; Length 334;  
 Best Local Similarity 50.4%; Pred. No. 2.8e-67;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 QY 10 DSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIOGFCGSLVEFVKN 69  
 Db 75 DESSESLTEGIRNSLIRQEDSIIFGLIERAKYCYNADTYDPTAFDWDGNGSLVEFVKN 134

QY 70 TEAIQAKAGRYKNEENAFPPENLPPSIVPSYSGFKQFLHPCAASININKSIWYKFKELL 129  
 Db 135 TEKLHAKVGRFKSPDBHPPFPDLPBPMLPFPLOYPKVLHFAADSIININKIWNMYFRDIV 194  
 QY 130 PLLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRDAPQDYEPLIRAKDKKGL 189  
 Db 195 PRLVKGGDGNYGSTAVCAICLQCSKRIHYGKFAEAKFQASPAYESAIAQDKRL 254  
 QY 190 MKLITFTSVETVRKVEKKAIVFGQVNLNSDDNENR-----KFDPSVASSLYK 241  
 Db 255 MDMLTFTPTVEDAIKXRVEMKTRTYGQEVKVGMBEKEEBEENSHVYKISPIVGLDLYG 314  
 QY 242 NWVPLTKVEQVYLLRLD 261  
 Db 315 DWIMPLTKVEQVYLLRLD 334

RESULT 4  
 US-09-454-279-14  
 Query Match 50.9%; Score 681.5; DB 4; Length 257;  
 Best Local Similarity 52.1%; Pred. No. 7e-66;  
 Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming  
 TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 FILE REFERENCE: BB1299 US NA  
 CURRENT APPLICATION NUMBER: US/09/454,279  
 CURRENT FILING DATE: 1999-12-03  
 EARLIER APPLICATION NUMBER: 60/110,845  
 EARLIER FILING DATE: 1998-12-04  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: Microsoft Office 97  
 SEQ ID NO 14  
 LENGTH: 257  
 TYPE: PRT  
 ORGANISM: Oryza sativa  
 US-09-454-279-14

Query Match 50.9%; Score 681.5; DB 4; Length 257;  
 Best Local Similarity 52.1%; Pred. No. 7e-66;  
 Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;  
 QY 6 EQSPDSGNVYTLASVREDLVROEDTIIYGLIERAKFPSNSHTYDEKVAQIOGFCGSLVEF 65  
 Db 1 KORIDQSEILTDNIRTSILVROEDSIIFSLIERAQFCYNADIYDKNAFHVDGFGSLVEF 60  
 QY 66 VVKNTAIAQAKAGRYKNEENAFPPENLPPSIVPSYSGFKQFLHPCAASININKSIWYK 125  
 Db 61 WURETEKLUHQVGRYKSPDEHPFEDLPPLPQYKVLHPIADSIININKSIWYK 120  
 QY 126 KELLPLATSGDGNAGTAANDLSLQSIERRHYGKFAEVRDAPQDYEPLIRAKD 185  
 Db 121 DELLPLRVKSGDGNYGSSALCDTICLQALSRIHYGKFAEAKFQESPEAYMPIAIOQ 180  
 QY 186 KEGMLKLLTFTSVETVRKVEKKAIVFGQVNLNSDDNENR--KFDPSVASSLYK 244  
 Db 181 CDOLMHLITETVETRAIEHRVEAKAKIFGQVLDIAGDNGAPPYKIRPSLVABEYSYRI 240  
 QY 245 IPLTKVEQVYLLRLD 261  
 Db 241 MELTKVEQVAVLLRLD 257

RESULT 5  
 US-09-454-279-6  
 Query Match 52.0%; Score 696; DB 4; Length 334;  
 Best Local Similarity 50.4%; Pred. No. 2.8e-67;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 GENERAL INFORMATION:  
 APPLICANT: Falco, S. Carl  
 APPLICANT: Famodu, Omolayo O.  
 APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 6

; LENGTH: 132

; TYPE: PRT

; ORGANISM: Glycine max

; US-09-454-279-6

Query Match 50.3%; Score 673; DB 4; Length 132;

Best Local Similarity 98.5%; Pred. No. 2.1e-65;

Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 DSGNVYTLASVREDLVQRQEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVKX 69

DB 1 DSGNVYTLASVREDLVQRQEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVKX 60

QY 70 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 129

DB 61 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 120

QY 130 PLLATSGDGGNY 141

DB 121 PLLATSGDGGNY 132

#### RESULT 6

US-09-454-279-12

; Sequence 12, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 12

; LENGTH: 331

; TYPE: PRT

; ORGANISM: Zea mays

; US-09-454-279-12

Query Match 50.1%; Score 670.5; DB 4; Length 331;

Best Local Similarity 51.4%; Pred. No. 1.6e-64;

Matches 130; Conservative 48; Mismatches 74; Indels 1; Gaps 1;

QY 10 DSGNVYTLASVREDLVQRQEDTIIYGLIERAKFPNSHTYDEKYAQIQGFCGSLVEFVVKX 69

DB 79 DRSDILTDISQVLINLEDSIIIFGLIERAQFCYNADTYDSNAFMDGFGGSLVEIVRE 138

QY 70 TEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELL 129

DB 139 TEKLHAQVGRYKSPDEHPFESKDLPEPLPPMOYPRVLHPHADSININKSIWKYFDELL 198

QY 130 PLLATSGDGGNYAQTAANDLSLQISRRRIHYGKFAEVKFRDAPODYEPLIRAKDEGL 189

DB 199 PRLVKESDGNAGSSALCDTTCIQLASRRRIHYGKFAEVKFRDAPODYEPLIRAKDEGL 258

QY 190 MKLLTFTSVETVRKVEKAVVFGQEVNLSDDNEN-RKPDPSVASSLYKQWVPLT 248

DB 259 MNLVTETVERAIEHRVEAKAFIFGQEVNIGAKDNGSPVYKIRPSLVAELSYRIMPLT 318

QY 249 KEVQVEYLLRLD 261

DB 319 KEVEYALLRLD 331

#### RESULT 7

US-09-454-279-18

; Sequence 18, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 18

; LENGTH: 207

; TYPE: PRT

; ORGANISM: Triticum aestivum

; US-09-454-279-18

Query Match 45.0%; Score 602; DB 4; Length 207;

Best Local Similarity 57.8%; Pred. No. 2.2e-57;

Matches 118; Conservative 27; Mismatches 53; Indels 6; Gaps 2;

QY 64 EFVVKNTAEIQAKAGRYKNPEENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKX 123

DB 4 EFFVREAEVLHAKAGHYQKPEDVPFSSQDLPPVPFTKGRPKVLHPFASLVCVNDALWKX 53

QY 124 YFKELLPLATSGDGGNYAQTAANDLSLQISRRRIHYGKFAEVKFRDAPODYEPLIRA 183

DB 64 YFNELLPLTADGDDGNYAETVALDFACLOALSRRRIHCGKYAEVYKFKDAPQDYSPPIRA 123

QY 184 KDKGKMLKLTFTSVETVRKVEKAVVFGQEVNLSDDNEN-RKPDPSVASSLYKQWVPLT 237

DB 124 KDTNALMNLFTAVEKVKVEKARIFQGNVTLEDSVKGQDQDADCSHCKVDPKVLS 183

QY 238 SLYKNWVPLTKVEQVYLLRLD 261

DB 184 KLYDMWVPLTKDVEYVLLRLD 207

#### RESULT 8

US-09-454-279-8

; Sequence 8, Application US/09454279

; Patent No. 6627798

; GENERAL INFORMATION:

; APPLICANT: Falco, S. Carl

; APPLICANT: Famodu, Omolayo O.

; APPLICANT: Lee, Jian-Ming

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES

; FILE REFERENCE: BB1299 US NA

; CURRENT APPLICATION NUMBER: US/09/454,279

; CURRENT FILING DATE: 1999-12-03

; EARLIER APPLICATION NUMBER: 60/110,845

; EARLIER FILING DATE: 1998-12-04

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 8

; LENGTH: 154

; TYPE: PRT

; ORGANISM: Triticum aestivum

; US-09-454-279-8

Query Match 36.5%; Score 489; DB 4; Length 154;

Best Local Similarity 59.7%; Pred. No. 2.8e-45;

Matches 92; Conservative 22; Mismatches 40; Indels 0; Gaps 0;



QY	64	EFVKNTEAIQAKAGRYKNPEENAFFPENPPSPVSPKQFLHPCAASINIKSWRM	123
Db	1	EFFVEAEVLHAKAGHYKQPDVFDFSQDLPPVPFTKGRPKVLHPFASIVCVNDAIWRM	60
QY	124	YFKEILLPLATSGDGNVQAOTANDLSLQISRRHYGKFVAEVKFRAPQDYEPLIRA	183
Db	61	YFNEILLPUFTADGDGNVTAETALDFCQLQALSRHICRGYVAEVKFKDAPQDYSPEIRA	120
QY	184	KDKGIMKLLTFTSYVEETVRKREKAVVFGQEV	217
Db	121	KDTNALMLLTFTVAEEKVKRVRKKKARIQGVN	154

```

RESULT 9
US-09-454-279-2
; Sequence 2, Application US/09454279
; Patent No. 662798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Ramodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454, 279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Zea mays
US-09-454-279-2

```

```

RESULT 10
US-09-454-279--4
; Sequence 4, Application US/09454279
; Patent No. 6627798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omoiyao O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (49)
; FEATURE:

```

```

RESULT 11
US-09-134-001C-4701
; Sequence 4701, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4701
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4701

```

RESULT 12  
US-09-543-681A-5022  
; Sequence 5022, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND  
; TITLE OF INVENTION: DIAGNOSTICS AND T  
; FILE REFERENCE: 2709.1002-001

;; CURRENT APPLICATION NUMBER: US/09/543,681A  
;; CURRENT FILING DATE: 2000-04-05  
;; PRIOR APPLICATION NUMBER: US 60/128,706  
;; PRIOR FILING DATE: 1999-04-09  
;; NUMBER OF SEQ ID NOS: 8344  
;; SEQ ID NO 5022  
;; LENGTH: 753  
;; TYPE: PRT  
;; ORGANISM: Proteus mirabilis  
US-09-543-681A-5022

Query Match 7.2%; Score 96; DB 4; Length 753;  
Best Local Similarity 19.0%; Pred. No. 0.2; Mismatches 86; Indels 110; Gaps 17;  
Matches 60; Conservative 60

QY 15 YTLASVREDLVROEDT-----IYGLIERAKFPNSHTYDE-----KYAQI---QGFCS 61  
Db 459 YSLAHFYADLLEEDVANNEEYAFSLYQARGVAYEATPHQATIKYKNIIDDQ---A 514  
QY 62 LVEFVVKNTAIOAKAGRYKNE-----ENAFPEKLPSPV- 99  
Db 515 LNIIVSELKDDLSRLTEKKRGEGVILFRYGMDSQYVIDFVRSYDDKIRKAIQPL 574  
QY 100 -----SYSPKQFLHPGAASININKSIWKYFKELLPLLATSGDDGNYAQTAAANDLS 150  
Db 575 LSQNTNLVFOYTIKNIHENSAN-NINEDNLKYYKDFKLA-----EIG 618  
QY 151 LLOSISRHHYGFVAEVRDAP-----QDY--BPLIRAKXKGL- 189  
Db 619 SIDAI-----KFIITKEYKDSAGDEIYQDYSDFKLTNITIKERLAWRKKADLGNFI 671  
QY 190 -MKLLT--FTSVETVRKVEKAVVFGQVNLNSDDNENRKPDPSPVASSLYKNWVIP 246  
Db 672 CLAKELAIYRDGEGVRKNI-AKASEYERRISNSHFLESEKYE-----IDRRW- 721  
QY 247 LTKVQ---VEYLLLR 259  
Db 722 --KEKRDNTLDYLLNQ 735

## RESULT 13

US-09-074-579-3

;; Sequence 3, Application US/09074579  
;; Patent No. 6001596  
;; GENERAL INFORMATION:  
;; APPLICANT: Hillman, Jennifer L.  
;; APPLICANT: Guegler, Karl J.  
;; APPLICANT: Patterson, Chandra  
;; TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE  
;; TITLE OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR  
;; NUMBER OF SEQUENCES: 5  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;; STREET: 3174 Porter Drive  
;; CITY: Palo Alto  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94304  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/074,579  
;; FILING DATE: HEREMITH  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Cerrone, Michael C  
;; REGISTRATION NUMBER: 39,132

;; REFERENCE/DOCKET NUMBER: PF-0505 US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-855-0555  
;; TELEFAX: 650-845-4166  
;; TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 946 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENEBANK

CLONE: g133985

US-09-074-579-3

Query Match 6.9%; Score 92; DB 3; Length 946;

Best Local Similarity 19.2%; Pred. No. 0.78;

Matches 56; Conservative 43; Mismatches 89; Indels 104; Gaps 13;

QY 19 SVREDLVROEDTIIYGLIERAKFPNSHTYDEKYAQIQGCSLVEFVVKNTAIOAKAG 78  
Db 263 SCRTAVDGLVLYD-VKR-----EKAGELEVFNGYVHF----- 298  
QY 79 RYKPEENAFPEENLPSPVPSYFKQFLHPGAASININKSIWKYFKELLPLLATSGDD 138  
Db 299 -----FAPDNLDP--IP-----KNILF-----VIDVSGSMGWGKKQTVKAMKTLDD 339  
QY 139 -----GNYAQTAAANDLSLOSISRHHYK-FVAEVRDAPQDYELIRA-- 183  
Db 340 LRAEDHFSVIDFNQIRWNRDLFQLQ--KHRLQIAKRYIEKIOPSGGTNINEALLRAIF 397  
QY 184 -----KKEGLMKULTTSVETVRKVEKAVVFGQVNLNS 221  
Db 398 ILNEANNLGLDLPNSVLIILVSDGPTVGLKLSKIQKNVKENIQDNISLFLSGMGDFV 457  
QY 222 DDN-----DNENRKFDPSPVASSLYKNW-----VIPTKEVQVEY 255  
Db 456 DYDFLKRLSNENH-----GIAQRIYGNQDTSSQLKFFYNQVSTPLLRNVQFNY 505

## RESULT 14

US-09-388-774-3

;; Sequence 3, Application US/09388774  
;; Patent No. 6228991  
;; GENERAL INFORMATION:  
;; APPLICANT: Hillman, Jennifer L.  
;; APPLICANT: Guegler, Karl J.  
;; APPLICANT: Patterson, Chandra  
;; TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE  
;; TITLE OF INVENTION: INHIBITOR HEAVY CHAIN PRECURSOR  
;; NUMBER OF SEQUENCES: 5  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
;; STREET: 3174 Porter Drive  
;; CITY: Palo Alto  
;; STATE: CA  
;; COUNTRY: USA  
;; ZIP: 94304  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/388,774  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 09/074,579  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Cerrone, Michael C

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

Query Match 6.9%; Score 92; DB 3; Length 946;  
Best Local Similarity 19.2%; Pred. No. 0.78;  
Matches 56; Conservative 43; Mismatches 89; Indels 104; Gaps 13;  
QY 19 SVREDLVREQDIIYGLIERAKPPNSHTYDEKVAIQGFCGSLVEFVVKNTFAIQAKA 78  
Db 263 SRETAVDGELVLYD-VKR-----BEKAGELEVFNQYFVHF----- 298  
QY 79 RYKNPEENAFPPENLPSPISVPSYFKQFLHPCMAASININKSIWKMYFKELLPPLATSGDD 138  
Db 299 -----FAPDNLDP--IP-----KNILF-----VIDVSGSMWGVQKQTVEMAKTILDD 339  
QY 139 -----GNVQAQTAANDLSLQISRIHYGK-FVAEVRFRDAPQDYEPILIRA-- 183  
Db 340 LRAEDHFSVIDFNQIRTWRLDFLQ--KRLQIAKYIEKIQPSGGTINEALLRAIF 397  
QY 184 -----KDEGLMKLLTFTSVETVRKVEKAVVFGO--EVNLSNDDNDNEN 228  
Db 398 ILNEANNLGLDPNSVSLIIVSDGDPVTGELKLSKIQKNVKNENIODNISLSLGMGFDV 457  
QY 222 DDN-----DNEKRFDPSSVASSLYKNW-----VIPLTKVQVVEY 255  
Db 458 DYDFLKRSLNENH-----GIAQRIYGNQDTSSQLKXFNQVSTPLLRNVQFNY 505

Search completed: September 25, 2004, 02:55:14  
Job time : 24 secs

Query Match 6.8%; Score 90.5; DB 3; Length 554;  
Best Local Similarity 22.2%; Pred. No. 0.49;  
Matches 51; Conservative 37; Mismatches 85; Indels 57; Gaps 11;  
US-09-319-989-6

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

Query Match 6.8%; Score 90.5; DB 3; Length 554;  
Best Local Similarity 22.2%; Pred. No. 0.49;  
Matches 51; Conservative 37; Mismatches 85; Indels 57; Gaps 11;  
US-09-319-989-6

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

Query Match 6.8%; Score 90.5; DB 3; Length 554;  
Best Local Similarity 22.2%; Pred. No. 0.49;  
Matches 51; Conservative 37; Mismatches 85; Indels 57; Gaps 11;  
US-09-319-989-6

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

Query Match 6.8%; Score 90.5; DB 3; Length 554;  
Best Local Similarity 22.2%; Pred. No. 0.49;  
Matches 51; Conservative 37; Mismatches 85; Indels 57; Gaps 11;  
US-09-319-989-6

REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0505 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 946 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENE BANK  
CLONE: gi33985  
US-09-388-774-3

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 02:54:21 ; Search time 73 Seconds  
(without alignments)  
1149.683 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTKVQVEYLLRRLD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTU5\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1339	100.0	261	9	US-09-454-279-16
2	1339	100.0	261	15	US-10-624-061-16
3	1339	100.0	297	12	US-10-425-114-48703
4	1315	98.2	261	12	US-10-424-599-235878
5	970	72.4	199	12	US-10-425-114-43606
6	818	61.1	265	9	US-09-454-279-22
7	818	61.1	265	15	US-10-624-061-22
8	696	52.0	334	9	US-09-454-279-21
9	685.5	51.2	313	16	US-10-624-061-21
10	685.5	50.9	257	9	US-10-437-963-196439
11	681.5	50.9	257	9	US-09-454-279-14
12	681.5	50.9	257	15	US-10-624-061-14
13	678.5	50.7	255	16	US-10-437-963-151855
14	673	50.3	132	9	US-09-454-279-6
15	673	50.3	132	15	US-10-624-061-6

16	671	50.1	264	12	US-10-425-114-66528	Sequence 66528, A
17	670.5	50.1	331	9	US-09-454-279-12	Sequence 12, Appl
18	670.5	50.1	331	15	US-10-624-061-12	Sequence 12, Appl
19	668.5	49.9	350	16	US-10-767-701-45917	Sequence 45917, A
20	661	48.6	320	12	US-10-425-114-68586	Sequence 68586, A
21	650	48.5	315	12	US-10-424-599-266862	Sequence 266862, A
22	636.5	47.5	313	12	US-10-425-114-65608	Sequence 65608, A
23	602	45.0	207	9	US-09-454-279-18	Sequence 18, Appl
24	602	45.0	207	15	US-10-624-061-18	Sequence 18, Appl
25	559	41.7	210	16	US-10-767-701-32934	Sequence 32934, A
26	543	40.6	316	16	US-10-437-963-186040	Sequence 186040, A
27	523	39.1	306	16	US-10-437-963-186047	Sequence 186047, A
28	515	38.5	489	16	US-10-437-963-103185	Sequence 103185, A
29	502.5	37.5	186	12	US-10-425-114-52677	Sequence 52677, A
30	489	36.5	154	9	US-09-454-279-8	Sequence 8, Appl
31	489	36.5	154	15	US-10-624-061-8	Sequence 8, Appl
32	443	33.1	109	12	US-10-424-599-224668	Sequence 224668, A
33	423	31.6	280	13	US-10-042-059A-2	Sequence 2, Appl
34	234.5	17.5	97	12	US-10-425-114-38176	Sequence 38176, A
35	188.5	14.1	83	12	US-10-425-114-52545	Sequence 52545, A
36	188.5	14.1	89	12	US-10-424-599-272410	Sequence 272410, A
37	147	11.0	93	9	US-09-454-279-2	Sequence 2, Appl
38	147	11.0	93	15	US-10-624-061-2	Sequence 2, Appl
39	140	10.5	118	12	US-10-424-599-276200	Sequence 276200, A
40	131	9.8	62	9	US-09-454-279-4	Sequence 4, Appl
41	131	9.8	62	15	US-10-624-061-4	Sequence 4, Appl
42	112	8.4	59	12	US-10-424-599-284420	Sequence 284420, A
43	103	7.7	781	9	US-09-815-242-11269	Sequence 11269, A
44	103	7.7	781	12	US-10-282-122A-58574	Sequence 58574, A
45	95.5	7.1	1556	15	US-10-369-493-1796	Sequence 1796, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-454-279-16  
; Sequence 16, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Pamodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BBI299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 16  
; LENGTH: 261  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-09-454-279-16

Query Match	100.0%;	Score 1339;	DB 9;	Length 261;
Best Local Similarity	100.0%;	Pred. No. 1.2e-125;		
Matches 261;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAKAAEQSPDSGNVYTLASVREDLVQRDIIIVGLIERAKFPNSHTYDEKYAIOGFCG	60	
Db	1	MAKAAEQSPDSGNVYTLASVREDLVQRDIIIVGLIERAKFPNSHTYDEKYAIOGFCG	60	
QY	61	SLVEFVVKNTAEIAQAKAGRYKKNPEENAFFPENPPSVPSYFKQFLHPGAASININKSI	120	
Db	61	SLVEFVVKNTAEIAQAKAGRYKKNPEENAFFPENPPSVPSYFKQFLHPGAASININKSI	120	
QY	121	WKMYFKELLPLLATSGDDGNYAQTAAANDLSLQSIISRIHYKGFVAEVKFRDPAQDYEPL	180	
Db	121	WKMYFKELLPLLATSGDDGNYAQTAAANDLSLQSIISRIHYKGFVAEVKFRDPAQDYEPL	180	

QY 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 DB 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 QY 241 KMWIPLTKVEQVEYLLRLD 261  
 DB 241 KMWIPLTKVEQVEYLLRLD 261

RESULT 2  
 US-10-624-061-16  
 ; Sequence 16, Application US/10624061  
 ; Publication No. US20040019929A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Famodu, Omolayo O.  
 ; APPLICANT: Lee, Jian-Ming  
 ; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 ; FILE REFERENCE: BB1299 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/624,061  
 ; CURRENT FILING DATE: 2003-07-21  
 ; PRIOR APPLICATION NUMBER: US/09/454,279  
 ; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 16  
 ; LENGTH: 261  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 US-10-624-061-16

Query Match 100.0%; Score 1339; DB 15; Length 261;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-125;  
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAKAAEQSPDSGNVYTLASVREDLVREDLVRQEDTIIYGLIERAKFPSNSHTYDEKYAQIOGFCG 60  
 DB 1 MAKAAEQSPDSGNVYTLASVREDLVREDLVRQEDTIIYGLIERAKFPSNSHTYDEKYAQIOGFCG 60  
 QY 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
 DB 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
 QY 121 WKMYFKELLPLLATSGDGNVYTAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYEPL 180  
 DB 121 WKMYFKELLPLLATSGDGNVYTAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYEPL 180  
 QY 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 DB 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 QY 241 KMWIPLTKVEQVEYLLRLD 261  
 DB 241 KMWIPLTKVEQVEYLLRLD 261

RESULT 3  
 US-10-425-114-48703  
 ; Sequence 48703, Application US/10425114  
 ; Publication No. US20040034898A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei  
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
 ; FILE REFERENCE: 38-21(53313/B  
 ; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 73128  
 ; SEQ ID NO 48703  
 ; LENGTH: 297  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: 700677130\_FLI.pep  
 US-10-425-114-48703

Query Match 100.0%; Score 1339; DB 12; Length 297;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-125;  
 Matches 261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAKAAEQSPDSGNVYTLASVREDLVREDLVRQEDTIIYGLIERAKFPSNSHTYDEKYAQIOGFCG 60  
 DB 37 MAKAAEQSPDSGNVYTLASVREDLVREDLVRQEDTIIYGLIERAKFPSNSHTYDEKYAQIOGFCG 96  
 QY 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
 DB 97 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 156  
 QY 121 WKMYFKELLPLLATSGDGNVYTAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYEPL 180  
 DB 157 WKMYFKELLPLLATSGDGNVYTAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYEPL 216  
 QY 181 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 240  
 DB 217 IRAKDEGLMKLLTFTSVETVRKVEKAVVFGQEVNLSDDNDNENRKFDPVASSLY 276  
 QY 241 KMWIPLTKVEQVEYLLRLD 261  
 DB 277 KMWIPLTKVEQVEYLLRLD 297

RESULT 4  
 US-10-424-599-235878  
 ; Sequence 235878, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
 ; FILE REFERENCE: 38-21(53223/B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 235878  
 ; LENGTH: 261  
 ; TYPE: PRT  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55025C.1.pep  
 US-10-424-599-235878

Query Match 98.2%; Score 1315; DB 12; Length 261;  
 Best Local Similarity 97.7%; Pred. No. 3.1e-123;  
 Matches 255; Conservative 6; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAKAAEQSPDSGNVYTLASVREDLVREDLVRQEDTIIYGLIERAKFPSNSHTYDEKYAQIOGFCG 60  
 DB 1 MAKAAEQSPDSGNVYTLASVREDLVREDLVRQEDTIIYGLIERAKFPSNSHTYDEKYAQIOGFCG 60  
 QY 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
 DB 61 SLVEFVVKNTETAIQAKAGRYKKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSI 120  
 QY 121 WKMYFKELLPLLATSGDGNVYTAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYEPL 180  
 DB 121 WKMYFKELLPLLATSGDGNVYTAQTAANDLSLQSISSRIHYGKFAVVKFRDAPQDYEPL 180

QY 181 IRAKDEGLMKLLTFTSVETVRKEKKAIVFGQEVNLSDDNDNENRKFDPSPVASSLY 240  
Db 181 IRAKDEGLMKLLTFTSVETVRKEKKAIVFGQEVNLSDDNDNENRKFDPSPVASSLY 240  
QY 241 KMWIPLTKVQVVEYLLRLD 261  
Db 241 KMWIPLTKVQVVEYLLRLD 261  
RESULT 5  
US-10-425-114-43606  
; Sequence 43606, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 43606  
; LENGTH: 199  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3050-017-E9\_FLI.pgp  
US-10-425-114-43606

Query Match 72.4%; Score 970; DB 12; Length 199;  
Best Local Similarity 94.0%; Pred. No. 8.2e-89;  
Matches 187; Conservative 8; Mismatches 4; Indels 0; Gaps 0;  
QY 63 VEFVKNTEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIWK 122  
Db 1 VEFVKNTEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININKSIWK 60  
QY 123 MYFELLPLLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDYELIR 182  
Db 61 MYFELLPLVATSGDDGNYAQTAAANDLSLQALSRRIHYGKFAVFKFDPADQDYELIR 120  
QY 183 AKDKEGLMKLLTFTSVETVRKEKKAIVFGQEVNLSDDNDNENRKFDPSPVASSLYKN 242  
Db 121 AKDKEGLMKLLTFTSVETVRKEKKAIVFGQEVNLSDDNDNENRKFDPSPVASSLYKN 180  
QY 243 WVIPLTKVQVVEYLLRLD 261  
Db 181 WVIPLTKVQVVEYLLRLD 199

RESULT 6  
US-09-454-279-22  
; Sequence 22, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454, 279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22

; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-454-279-22  
Query Match 61.1%; Score 818; DB 9; Length 265;  
Best Local Similarity 61.1%; Pred. No. 2.1e-73;  
Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
QY 1 MAKAAEQSPDSG--NVYTLASVREDIVRQEDTIYGLIERAKFPNSHNTYDEKYAQIQGF 58  
Db 1 MARVFESDSGSGCNSVLSDLIRESLIRQEDTIYGLIERAKFPNSPAFESRCLDSGS 60  
QY 59 CGSLVEFVKNTEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININK 118  
Db 61 FSSLTEFFVRETEIIQAKGRYEPENPFLENIPHSVPTHKYPSPALHPKALSININK 120  
QY 119 SIWKMYFKEKLLPLLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDY 178  
Db 121 QIWDIYFKEKLLPLFVKPGDDGNYPTAASDLACLQALSRRIHYGKFAVFKFDPADQDY 180  
QY 179 PLIRAKDEGLMKLLTFTSVETVRKEKKAIVFGQEVNLSDDNDNENRKF--DPSVA 236  
Db 181 PAIRAQDREALMKLLTTFEKEEMVKRVQKKAETFGQEVKNSGYGDSKKYKVDPLLA 240  
QY 237 SSLYKNWVPLTKVQVVEYLLRLD 261  
Db 241 SRIYGEWVPLTKVQVVEYLLRLD 265

RESULT 7  
US-10-624-061-22  
; Sequence 22, Application US/10624061  
; Publication No. US20040019929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 22  
; LENGTH: 265  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-624-061-22

Query Match 61.1%; Score 818; DB 15; Length 265;  
Best Local Similarity 61.1%; Pred. No. 2.1e-73;  
Matches 162; Conservative 31; Mismatches 68; Indels 4; Gaps 2;  
QY 1 MAKAAEQSPDSG--NVYTLASVREDIVRQEDTIYGLIERAKFPNSHNTYDEKYAQIQGF 58  
Db 1 MARVFESDSGSGCNSVLSDLIRESLIRQEDTIYGLIERAKFPNSPAFESRCLDSGS 60  
QY 59 CGSLVEFVKNTEAIQAKAGRYKNPEENAFPPENLPPSIVPSYFKOFLHPGAASININK 118  
Db 61 FSSLTEFFVRETEIIQAKGRYEPENPFLENIPHSVPTHKYPSPALHPKALSININK 120  
QY 119 SIWKMYFKEKLLPLLATSGDDGNYAQTAAANDLSLQSTSRRIHYGKFAVFKFDPADQDY 178  
Db 121 QIWDIYFKEKLLPLFVKPGDDGNYPTAASDLACLQALSRRIHYGKFAVFKFDPADQDY 180  
QY 179 PLIRAKDEGLMKLLTFTSVETVRKEKKAIVFGQEVNLSDDNDNENRKF--DPSVA 236  
Db 181 PAIRAQDREALMKLLTTFEKEEMVKRVQKKAETFGQEVKNSGYGDSKKYKVDPLLA 240





Db 232 ITAQDCDQMLHLLTYETVERAIEHRVAKAFQGVDELGAEDNGAPPYKIRPSLVAEI 291  
QY 240 YKQWVPLTKVEQVVEYLLRLD 261  
Db 292 YSYRIMPLTKVEQVAYLLRLD 313

RESULT 11  
US-09-454-279-14  
; Sequence 14, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 14  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-09-454-279-14

Query Match 50.9%; Score 681.5; DB 9; Length 257;  
Best Local Similarity 52.1%; Pred. No. 1e-59;  
Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 6 EQSPDSGNVYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKYAQIQGFCGSLVEF 65  
Db 1 KORIDQSEILTDNIRTSVLRQEDSIIFSLERAQFCYNADIYKNAFHVDFGSLVEF 60  
QY 66 VVKNTAIOAKAGRYKNPEENAFPPENLPPSVPSYFKQFLHPGAASININKSIWKYF 125  
Db 61 MVRETEKLHQGVGRYKSPDEHFFPDLPEPLLPLOYKVLHPADSIINKEIKWYF 120  
QY 126 KELLPLLATSGDGNVYAQTAAANDLSLQISRRHYGKFAVVKFRDAPQDYEPILRAKD 185  
Db 121 DELLPLRVKESDNGYSSALCDTICLQALSRIHYGKFAVAKFQESPEAYMPAIIAQD 180  
QY 186 KEGMLKLLFTSVETVRKVEKKAUVFGQEVNLSDDNDNENR-KFDPSPVASSLYKNWV 244  
Db 181 CDQLMHLTYETVERAIEHRVAKAFQGVDELGAEDNGAPPYKIRPSLVAEIYSYRI 240  
QY 245 IPLTKEVQVEYLLRLD 261  
Db 241 MPLTKEVQVAYLLRLD 257

RESULT 12  
US-10-624-061-14  
; Sequence 14, Application US/10624061  
; Publication No. US20040019929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624,061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 60/110,845  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 14  
; LENGTH: 257  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-624-061-14

Query Match 50.9%; Score 681.5; DB 15; Length 257;  
Best Local Similarity 52.1%; Pred. No. 1e-59;  
Matches 134; Conservative 44; Mismatches 78; Indels 1; Gaps 1;

QY 6 EQSPDSGNVYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKYAQIQGFCGSLVEF 65  
Db 1 KORIDQSEILTDNIRTSVLRQEDSIIFSLERAQFCYNADIYKNAFHVDFGSLVEF 60  
QY 66 VVKNTAIOAKAGRYKNPEENAFPPENLPPSVPSYFKQFLHPGAASININKSIWKYF 125  
Db 61 MVRETEKLHQGVGRYKSPDEHFFPDLPEPLLPLOYKVLHPADSIINKEIKWYF 120  
QY 126 KELLPLLATSGDGNVYAQTAAANDLSLQISRRHYGKFAVVKFRDAPQDYEPILRAKD 185  
Db 121 DELLPLRVKESDNGYSSALCDTICLQALSRIHYGKFAVAKFQESPEAYMPAIIAQD 180  
QY 186 KEGMLKLLFTSVETVRKVEKKAUVFGQEVNLSDDNDNENR-KFDPSPVASSLYKNWV 244  
Db 181 CDQLMHLTYETVERAIEHRVAKAFQGVDELGAEDNGAPPYKIRPSLVAEIYSYRI 240  
QY 245 IPLTKEVQVEYLLRLD 261  
Db 241 MPLTKEVQVAYLLRLD 257

RESULT 13  
US-10-437-963-151855  
; Sequence 151855, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 151855  
; LENGTH: 255  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_51960C.1.pap  
US-10-437-963-151855

Query Match 50.7%; Score 678.5; DB 16; Length 255;  
Best Local Similarity 53.1%; Pred. No. 2e-59;  
Matches 135; Conservative 43; Mismatches 63; Indels 13; Gaps 3;

QY 16 TLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKYAQIQGFCG--SLVEFVVKNTTEAI 73  
Db 7 SLAAVRDALVREDSIVFALIERARRPRNAPAYAAAAA---AGGSLAEFFVREAEVL 61  
QY 74 QAKAGRYKNPEENAFPPENLPPSVPSYFKQFLHPGAASININKSIWKYFKELLPLLA 133  
Db 62 HAKAGYQKEDVFFFPQDLPSFLPTKPKVLHSPASVSVDALWKYFNELLPLFT 121  
QY 134 TSGDDGNVYAQTAAANDLSLQISRRHYGKFAVVKFRDAPQDYEPILRAKDEGLMKLL 193  
Db 122 VDGDDGNVYAETVALDFACLKALSRRRIHGKYVAEVKFKDASQDYSPLIRAKDKALMNL 181

[illegible]

```

RESULT 14
US-09-454-279-6
; Sequence 6, Application US/09454279
; Publication NO. US20020184658A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Ramodu, Omolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1996-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Glycine max
US-09-454-279-6

```

Query Match	50.3%;	Score	673;	DB	9;	Length	132;
Best Local Similarity	98.5%;	Pred. No.	2.7e-59;				
Matches	130;	Conservative	0;	Mismatches	2;	Indels	0;
Gaps							
QY	10	DSGNVYTLASVREDLVROEDTIIYGLTERAKFPSNSHTYDEKYAQIQGCGSLVEFVVKV	69				
DB	1	DSGNVYTLASVREDLVROEDTIIYGLTERAKFPSNSHTYDEKYAQIQGCGSLVEFVVKV	60				
QY	70	TEAIOAKAGRYKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSIWMYFKELL	129				
DB	61	TEAIOAKAGRYKNPEENAFPPENLPSPISVPSYFQFLHPGAASININKSSGMYFKELL	120				
QY	130	PLLATSGDDGNY	141				
DB	121	PLLATSGDDGNY	132				

RESULT 15  
US-10-624-061-6  
; Sequence 6, Application US/10624061  
; Publication No. US2004001929A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BBI299 US NA  
; CURRENT APPLICATION NUMBER: US/10/624.061  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 6  
; LENGTH: 132  
; TYPE: PRT  
; ORGANISM: Glycine max  
US-10-624-061-6

```

Query Match      50.38; Score 673; DB 15; Length 132;
Best Local Similarity 98.5%; Pred No. 2,7e-59;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 DSGNYTLASVREDLVQSBTDIIYGLIERAKPFSNSHTYDERKACIOGFCGSLVEFVVKN 69
Db 1 DSGNYTLASVREDLVQSBTDIIYGLIERAKPFSNSHTYDERKACIOGFCGSLVEFVVKN 60

QY 70 TEAIOAKAGRYKYPEENAFPEPNLPSSIVPSYSFQKQFLHPGAASININKSITWKMYFKELL 129
Db 61 TEAIOAKAGRYKYPEENAFPEPNLPSSIVPSYSFQKQFLHPGAASININKSITWKMYFKELL 120

QY 130 PILATSGDGGNY 141
Db 121 PILATSGDGGNY 132

Search completed: September 25, 2004, 03:00:23
Job time : 74 secs

```

Search completed: September 25, 2004, 03:00:23  
Job time : 74 secs

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 07-Mar-2003



```
RESULT 5
T14902
Chorismate mutase (EC 5.4.99.5) 2, cytosolic - parsley (fragment)
C:Species: Petroselinum crispum (parsley)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 07-Mar-2003
C:Accession: T14902
R:Batiz, O.; Logemann, E.; Reinold, S.; Hahlbrock, K.
submitted to the EMBL Data Library, July 1997
A:Description: Extensive reprogramming of cellular metabolism by fungal elicitor or infection
A:Reference number: Z18251
A:Accession: T14902
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-82 <BAT>
A:Cross-references: EMBL:AF012867; NID:g2352930; PID:g2352931
A:Gene: CCM2
C:Function:
A:Description: intramolecular transferase; isomerase
A:Pathway: aromatic amino acid biosynthesis
C:Superfamily: chorismate mutase of the AroQ class, eukaryotic type
C:Keywords: aromatic amino acid biosynthesis; intramolecular transferase; isomerase; shikimate
Query Match 21.1%; Score 283; DB 2; Length 82;
Best Local Similarity 63.4%; Pred. No. 3.4e-16;
Matches 52; Conservative 10; Mismatches 20; Indels 0; Gaps 0;
QY 88 FFPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELLPPLATSGDGNVYAQTAAAN 147
DB 1 FPDNLPSSLLPQYNHPQLLHTPTAAININEIWDAVYNQLPLFTTGGDGNVYPTATS 60
QY 148 DLSLOSISRRHYGKFAEVK 169
DB 61 DLQLOAISRRHYGKFAEAK 82
RESULT 6
T14901
Chorismate mutase (EC 5.4.99.5) 1, chloroplast - parsley (fragment)
C:Species: Petroselinum crispum (parsley)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 07-Mar-2003
C:Accession: T14901
R:Batiz, O.; Logemann, E.; Reinold, S.; Hahlbrock, K.
submitted to the EMBL Data Library, July 1997
A:Description: Extensive reprogramming of cellular metabolism by fungal elicitor or infection
A:Reference number: Z18251
A:Accession: T14901
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-70 <BAT>
A:Cross-references: EMBL:AF012866; NID:g2352928; PID:g2352929
C:Genetics:
A:Gene: pCMI
A:Genome: nuclear
C:Function:
A:Description: intramolecular transferase; isomerase
A:Pathway: aromatic amino acid biosynthesis
C:Superfamily: chorismate mutase of the AroQ class, eukaryotic type
C:Keywords: aromatic amino acid biosynthesis; chloroplast; intramolecular transferase; shikimate
Query Match 14.0%; Score 188; DB 2; Length 70;
Best Local Similarity 51.4%; Pred. No. 1.3e-08;
Matches 36; Conservative 8; Mismatches 26; Indels 0; Gaps 0;
QY 90 PENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELLPPLATSGDGNVYAQTAAANDL 149
DB 1 PEHLPEPISPLQYQVLPVHTESININSKINDMYFRNLVPRLVKEDGNGYATVCDT 60
QY 150 SLIQSISRR 159
DB 61 ICLQALSKRI 70
```

```
RESULT 7
G95041
hypochemical protein SP0355 [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: G95041
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide-
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzap-
ple, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete genome sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: G95041
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74528.1; PID:g14971828; GSPDB:GN00164; TIGR:SP45
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0355
Query Match 7.8%; Score 105; DB 2; Length 360;
Best Local Similarity 24.7%; Pred. No. 0.64;
Matches 55; Conservative 34; Mismatches 76; Indels 58; Gaps 11;
QY 31 IYGLIERAKF-----PSN-SHTYDEKYAQIQGFCGSLVEFVVXNTEATQAKAGRYKNP 83
DB 26 IVYDLLESKGVESFYNPSPDFQYLYKEVRQKQ-----SFTKKQADAILYIPGGYFGE 78
QY 84 EENAFPPENLPPSIVPSYFKQFLHPGAASININKSIWKYFKELLPPLATSGDGNVYAQ 143
DB 79 GHNARFRDNL-----IQKRFPLDQ-----IWASYFKKPIGVILGAGPN----- 118
QY 144 TAANDLSLOSISRRHYGKFAEVKFRD-----AP-QDYEPLIRAKDEGLM 190
DB 119 --NDSLMNYGKRIINHAQFIT-VRDRSFDLSKLHSPSPVHETFDLISSKLREKT 174
QY 191 KLLFTSVETVRKV-----EKKAV-VFGQEVNLSDDNDN 226
DB 175 EQQLQKREAKDKIILVHYHNSKKALEKFAESISLFLNNPN 217
RESULT 8
B64135
Ribonucleoside-diphosphate reductase (EC 1.17.4.1) alpha chain - Haemophilus influenzae
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 18-Jun-1999
C:Accession: B64135
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.;
D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, J.
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64135
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-781 <TIGR>
A:Cross-references: GB:U32839; GB:I42023; NID:g1574508; PIDN:AAC23305.1; PID:g1574509; TJ
C:Genetics:
A:Start codon: GTG
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase large chain
C:Keywords: DNA replication; iron-sulfur protein; oxidoreductase
Query Match 7.7%; Score 103; DB 2; Length 781;
Best Local Similarity 22.9%; Pred. No. 2.7;
Matches 53; Conservative 36; Mismatches 104; Indels 38; Gaps 9;
QY 49 DEKVAIQGFCGSLVEFVKV-----NTEATQAKAGRYKNPEENAFPPENLPPSIVPSYS 102
```



A;Residues: 1-295 <EKI>  
A;Cross-references: EMBL:D44597; NID:g871938; PID:d1008597; PID:g871940  
C;Genetics:  
A;Cross-references: SGD:S0001937  
A;Map position: 6R  
A;Note: YPR041c  
C;Superfamily: yeast membrane protein YPR041c; dnaJ amino-terminal homology  
C;Keywords: transmembrane protein  
F;8-24/Domain: transmembrane #status predicted <TM1>  
F;44-108/Domain: dnaJ amino-terminal homology <DNJ>  
F;133-149/Domain: transmembrane #status predicted <TM2>

Query Match 7.0%; Score 93.5; DB 2; Length 295;  
Best Local Similarity 22.1%; Pred. No. 4.2;  
Matches 56; Conservative 42; Mismatches 96; Indels 59; Gaps 13;

Qy 15 YTLASVEDLVQRDDTIYGLIERAKFPSSHTDEKYAQIQGCGSLVFWKNTAEIQ 74  
Db :  
19 YAFTTIEIFQLQNEI-----STKYGDWNFY--KFLKLPLQNLSSTKEITQLRKLS 70  
Qy 75 AKGRYNKPENAFPE-NLPSPISVPYS----FKQFLHPCAAASININKSIWKMYFKELL 129  
Db :  
71 KKYPDKNPKRYKLYERLNATQLSNSSNRKIYDYLLONGFPNYDPHKG--GFYFSRMK 128  
Qy 130 P---LLATSGDDGNYAQTANDLSLLQTSRRRIHYGVFAEVKFRDAPQDIYPLIRAXD 185  
Db :  
129 PKTWFLAFAFTVVVNICQYI---ISIIQYRSORSRIENFISOCKQQD-----D 173  
Qy 186 KEGFLMKLLFTSVEEYVRKRVEKKAVVF-----GQEVNLNSDDNDNENRKFD 233  
Db :  
174 TNGLGVRQLTFKQHEKD-----EGKSLVRFSDVVVVEPDGSETLISPDLTK-----P 222  
Qy 234 SVASSLYKNWVIP 246  
Db :  
223 SVKNCLF--WRIP 233

RESULT 12  
D75014  
Hypothetical protein PAB1257 - Pyrococcus abyssi (strain Orsay)  
C;Species: Pyrococcus abyssi  
C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999  
C;Accession: D75014  
R;anonymous, Genoscope  
submitted to the EMBL Data Library, July 1999  
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru  
A;Reference number: A75001  
A;Accession: D75014  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1134 <XAW>  
A;Cross-references: GB:AJ248289; GB:AL096836; NID:g5458960; PIDN:CAB50554.1; PID:e151645  
A;Experimental source: strain Orsay  
C;Genetics:  
A;Gene: PAB1257

	Query Match	7.0%; Score 93.5; DB 2; Length 1134;	
	Best local Similarity	22.8%; Pred. No. 26;	
	Matches	47; Conservative 41; Mismatches 71; Indels 47; Gaps 9;	
Qy	18 ASVREDLVROEDTTIYGLIERAKFPSSNSHYDEKYAQIQGFCGSLVEFVKNTAEAIQAKA	77	
Dd	654 SEILKDLIVREKISEIIIE-----LYNQKIWCYSRHSFV-----EKITGKA	697	
Qy	78 GRYNKPENEAFFPENLPSPISVPYSFKQF--LH-----PGAAISININKSIWKMYFKELL	129	
Dd	698 RVYKDCECNA---EDRI SNLLNNLGPYYTMLHSNMGIVEGVGEAINIVSTIGEIVSKE-L	753	
Qy	130 PLLATSGDDGNYAQTAANDLSL-LQSISRRIHYCKFVAEVKFDPADQDIYEPLTRAKDKEG	188	
Dd	754 SRGTNIENVKTLDNRIGLDJDISLEDIAKK-----SRDYNEFLHSVEETI	798	
Qy	i89 LMKLLLTFTSV---EETVRKEVEKKAV	211	

```

|::||:|:::||:|
799 LRELLAYKGISTREHLKSEIDKLVS 824

Db

RESULT 13
A71683
hypothetical protein RP278 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Cession: A71683
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: A71683
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-371 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAAL4739.1; PID:g386083.
A:Experimental source: Strain Madrid E
C:Genetics:
A:Gene: RP278
```

```

Query Match      6.9%; Score 92; DB 2; Length 371;
Best Local Similarity 22.6%; Pred. No. 7.5;
Matches 64; Conservative 34; Mismatches 91; Indels 94; Gaps 14;

Qy      14 VYTLASVREDLVROBDTIYGLI-----ERAKF-----41
Db      117 LYSLLTKETDLIKANNIPESILTPDRLRLKCNLTPEQEERKKFEVLSILGSIIDDT 176
Qy      42 -PSNSHYDEKVAIQCGSLV--EFVQNTAEATCAKAGRYKNPEENAFPPENLPPSIV 98
Db      177 KKSNEH-YDKRANEINEQNKTIINEEPRKLUDEQKKIEGKKTQASEA-----LWKLLK 230
Qy      99 PSYSFKQF-----LHPGAASININKSIWMYFKELLPLPLATSGDDGNYAQATA 146
Db      231 PIY--KOMDEEREELFKLAELILPOYAQANIDKHA-KLYAKQYQTKIE---NDPNY----- 279
Qy      147 NDLISLQGISRIIHVGFKAVEKFPDAPQDYELIRAKDKEGMLKLLTTSVEETVKRV 206
Db      280 KELEKLEIVSKIEFYSK---KSKTND-----IINPNSTVVRULLKEKIDKESKKE 329
Qy      207 EKKAV-VFGQEVNLNSDDNDNENRKFDPSVASSLYKNWVPIPT 248
Db      330 MTEAINVIKQEVNLT-----KPFITPIPT 353

```

RESULT 14

S64410  
probable membrane protein YGR103w - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: hypothetical protein G5933  
C;Species: Saccharomyces cerevisiae  
C;Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 19-Apr-2002  
C;Accession: S64410, S64408  
R;Wedler, H.; Sharfe, M.; Wedler, E.; Wambutt, R.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64356  
A;Accession: S64410  
A;Molecule type: DNA  
A;Residues: 1-605 <WED>  
A;C;Cross-references: EMBL:Z72888; NID:gi323158; PID:gi323159; MIPS:YGR103w  
A;Experimental source: strain S288C  
R;Hernandez, K.; Weber, N.; Wipfli, P.; Schmidheini, T.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64392  
A;Accession: S64408  
A;Molecule type: DNA  
A;Residues: 1-18 <HER>  
A;C;Cross-references: EMBL:Z72888; MIPS:YGR103w  
A;Experimental source: strain S288C  
C;Genetics:  
A;Gene: SGD:RRP13

A:Cross-references: SGD:S0003335  
A:Map position: 7R  
C:Keywords: transmembrane protein  
F:217-233/Domain: transmembrane #status predicted <TMM>

Query Match 6.9%; Score 92; DB 2; Length 605;  
Best Local Similarity 20.5%; Pred. No. 15;  
Matches 62; Conservative 50; Mismatches 116; Indels 74; Gaps 13;

QY 2 AKAAQSPDGNVYTLASVRE-----DLVRQEDTIYGLIERAKFPFSNSHTYDEKYAQ 54  
DB 102 AKRLEENRDS---YTLDHIIKERYSPFDAIRDIDDALNMLFLFSLNPSTNQVSSKIIND 158  
QY 55 IQGFCGSLVEFVVKX-----TEATCAKAGRYKNPEENAFPPENLPPSI-- 97  
DB 159 AQKICNQWLAYAKERLVKRVFVSIKGVYQANIKGEEVRLVPFK---PPENIPSDVDF 215  
QY 98 VPSYGFQKFLHPGAASINIKSIWYFKELLPLLATSGDDGNVYAQTAANDLSLLQSI 157  
DB 216 RIMLTFLEFY-----STLHFVLYKLY-----TDSGLIYPPKLDLKKDKIIS- 257  
QY 158 RIHYGKFAVAFKRDAPQDYPL-----TRAKQKGLMKLLTF-----TSVEETVR 203  
DB 258 --GLSSYILSRQEDSLKLDPTETIEDVKVESLDASTKSAALNADAEANTDETEKEE 315  
QY 204 KEVEK-KAVVFGQEVNLSDDNDNENR-----KPDPSVASSLYKNWVIPLTKVQVE 254  
DB 316 KQKEKQKEQNEETELDTFEDNNKNGDILLIQSKYDSPA-SLFSAFVYVVSREVPID 374  
QY 255 YL 256  
DB 375 IL 376

Search completed: September 25, 2004, 02:54:47  
Job time : 27 secs

QY 176 DYEPLIRAKDKGLMKLLTF-----TSVEETVRKRVKKAUVFQGVNLSN----- 220  
DB 202 KOEMLSVLSVRRERLEKALSFMEABISVLOVEKRIRSRVRQMEKTQREYLLNEQMKAIQK 261  
QY 221 ----SDNDNENRKFPDPSVASSLYKNWVIPLTKVQVEYL--LRRL 260  
DB 262 ELGDSGSDGRDEVAEIEERITKT-----KLSKEAREKALAEELKKL 300

RESULT 15  
AF3361  
endopeptidase La (EC 3.4.21.53) [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 02-Jun-2003  
C:Accession: AF3361  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Eizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AF3361  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-823 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL52057.1; PID:G17982825; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0876  
A:Map position: 1  
C:Superfamily: ATP-dependent Lon protease  
C:Keywords: hydrolase; serine proteinase

Query Match 6.9%; Score 92; DB 2; Length 823;  
Best Local Similarity 22.0%; Pred. No. 22;  
Matches 63; Conservative 44; Mismatches 85; Indels 94; Gaps 14;

QY 4 RAEQSPDGNVY---TLASVREDLVRQEDTIYGLIE---RAKFP--SNSHTYDEKYAQI 55  
DB 80 AADDOPAPDAIYEIGTIANVLO-LKLPLDGVTKVVEGTAKAKISKFTDREDYHEAYA-- 136  
QY 56 QGFCGSLVEFVVKNTETAIQAKAGRYKNPEENAFPPENLPPSIYPSYFQKFLHPGAASIN 115  
DB 137 -----AALQ-----EPEDAVEIEALARSVVS--DFENY-----VK 165  
QY 116 INKSIWKYFKELLPLLATSGDDGNVYAQTAANDLSLLQSIERRIHYGKFAVAFKRDAPQ 175  
DB 166 LNKKI-----SPEVVGTSQIDDYSKLADTVASHLAI-----KTFE 201



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 25, 2004, 01:26:35 ; Search time 18 seconds

(without alignments)  
755.017 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYLASV.....NWVPLTKVEQVVEYLLRLD 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	698	52.1	334	1 CHMU_ARATH	P42738 arabidopsis
2	498.5	37.2	256	1 CHMU_YEAST	P32178 saccharomyc
3	462.5	34.5	251	1 CHMU_SCHPO	O13739 schizosacch
4	103	7.7	756	1 RIRI_HAEIN	P43754 haemophilus
5	101	7.5	737	1 SUV3_SACDO	O74727 saccharomyc
6	94.5	7.1	761	1 RIRI_BUCBP	Q89484 buchnera ap
7	93.5	7.0	295	1 YFLI_YEAST	P43613 saccharomyc
8	93	6.9	812	1 LON_BRUAB	O52605 bruceella ab
9	92	6.9	605	1 PESC_YEAST	P53261 saccharomyc
10	92	6.9	812	1 LON_BRUME	O8VHC6 bruceella me
11	92	6.8	812	1 LON_BRUSU	O8G017 bruceella su
12	91.5	6.8	510	1 ATPA_BUCAP	O51874 buchnera ap
13	91.5	6.8	631	1 THIC_SALTI	Q82326 salmonella
14	91.5	6.8	832	1 ALP6_SCHPO	Q9UEQ2 schizosacch
15	91	6.8	439	1 HSLU_CAMJE	O9PH10 campylobact
16	90.5	6.8	554	1 HAP4_YEAST	P14064 saccharomyc
17	90	6.7	693	1 EFG_STEMU	O8DVV4 streptococc
18	90	6.7	946	1 ITH2_MESAU	P97279 mesocricetu
19	88.5	6.6	1468	1 RPOE_AQUAE	O67762 aquifex aso
20	88	6.6	737	1 SUV3_YEAST	P32380 saccharomyc
21	87.5	6.5	631	1 THIC_SALTY	Q91317 salmonella
22	87.5	6.5	789	1 CAD9_HUMAN	Q9ULB4 homo sapien
23	87.5	6.5	2869	1 RBPI_PLAVB	Q00798 plasmodium
24	87	6.5	795	1 CDB5_HUMAN	Q9V5E4 homo sapien
25	87	6.5	936	1 VG2K_YEAST	P53253 saccharomyc
26	86.5	6.5	1653	1 CLH_YEAST	P22137 saccharomyc
27	86	6.4	946	1 ITH2_MOUSE	Q61703 mus musculus
28	85.5	6.4	918	1 IMB2_YEAST	P38217 saccharomyc
29	85	6.3	299	1 COMQ_BACSU	P33690 bacillus su
30	85	6.3	381	1 ARGJ_THETH	P96137 thermus th
31	85	6.3	601	1 SG2_RANRI	P30945 rana ridibu
32	84	6.3	1051	1 VP2_AHSV6	O71024 african hor
33	83.5	6.2	373	1 YIA6_YEAST	P40556 saccharomyc

## RESULT 1

ID	CHMU_ARATH	STANDARD;	PRT;	334 AA.
AC	P42738; Q9LS75; Q9SUJ5;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Chorismate mutase, chloroplast precursor [EC 5.4.99.5] (CM-1).			
GN	CM1 OR A13G29200 OR M0211.4.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsids.			
OX	NCBI taxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=940339841; PubMed=8224252;			
RA	Eberhard J., Raesecke H.-R., Schmid J., Amrhein N.;			
RT	"Cloning and expression in yeast of a higher plant chorismate mutase.			
RT	Molecular cloning, sequencing of the cDNA and characterization of the			
RT	Arabidopsis thaliana enzyme expressed in yeast.";			
RL	FEBS Lett. 334:233-236(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Landsberg erecta;			
RA	Kuhn R., Vogt E., Schmid J., Amrhein N., Schaller A.;			
RT	"Expression analysis of Arabidopsis thaliana genes for plastidic (CM1)			
RT	and cytosolic (CM2) chorismate mutases";			
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=20277480; PubMed=10819329;			
RA	Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;			
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. I.			
RT	Sequence features of the regions of 4,504,864 bp covered by sixty P1			
RL	and TAC clones.";			
CC	DNA Res. 7:131-135(2000).			
CC	!- CATALYTIC ACTIVITY: Chorismate = prephenate.			
CC	!- ENZYME REGULATION: Allosterically inhibited by tyrosine and			
CC	phenylalanine. Activated by tryptophan.			
CC	!- PATHWAY: BRANCH POINT OF THE BIOSYNTHETIC PATHWAY LEADING TO THE			
CC	THREE AROMATIC AMINO ACIDS, PHENYLALANINE, TYROSINE, & TRYPTOPHAN			
CC	(THE SHIKIMATE PATHWAY).			
CC	!- SUBCELLULAR LOCATION: Chloroplast.			
CC	!- SIMILARITY: Contains 1 chorismate mutase domain.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; 226519; CAA81286.1; -.			

DR EMBL; AJ242647; CAB54518.1; -;  
 DR EMBL; AB026657; BAB01816.1; -;  
 DR PIR; S38958; S38958.  
 DR HSP; P32178; 5CSM.  
 DR InterPro; IPR008238; Chor mut AroQ eu.  
 DR InterPro; IPR008951; ChorisMat mut II.  
 DR InterPro; IPR002701; ChorisMat mut.  
 DR Pfam; PF01817; ChorisMat mut.  
 DR PIRSF; PIRSF017318; Chor mut AroQ eu; 1.  
 KW Aromatic amino acid biosynthesis; isomerase; Chloroplast;  
 KW Transit peptide; Allosteric enzyme.  
 FT TRANSIT 1 60 CHLOROPLAST (POTENTIAL).  
 FT CHAIN 61 334 CHORISMATE MUTASE.  
 FT CONFLICT 253 253 A -> R (IN REF. 1).  
 FT CONFLICT 334 AA; 37594 MW; 4D477808E349B08C CRC64;  
 SQ SEQUENCE 334 AA; 37594 MW; 4D477808E349B08C CRC64;  
 Query Match 52.1%; Score 698; DB 1; Length 334;  
 Best Local Similarity 50.4%; Pred. NO. 2.8e-47;  
 Matches 131; Conservative 46; Mismatches 75; Indels 8; Gaps 1;  
 QY 10 DSGNVYTLASVREDLAVROEDRIIVGLIERAFPSNSHTYDEKYAQIOGFCGSLVEFVVKX 69  
 Db 75 DESESLTLEGINSIREDGSIIFLLERAKYCNADTYDTAFDMGFGSLVEYVWKG 134  
 QY 70 TEAIQAKAGRYKXNPENAFPPENIPPSIVPSYKQFLHPGAASININKSIWQYFKELL 129  
 Db 135 TEKLHAKVGRPKSPDEHPFFPDLLPEPMLPPLQVPKVLHFAADSININKKIWNMYFRDLV 194  
 QY 130 PLLATSGDGNVYAQTAANDLSLQISIRIHYGKVAEVRDAPQDYELIRAKDEGL 189  
 Db 195 PLVVKGGDGNVYGTAVCDALCLQCLSRIRHYGKVAEVRDAPQDYELIRAKDEGL 254  
 QY 190 MKLLFTTSVEETVRKVRKKAQVFCQVNLNSDNDNENR-----KDPSPVASSILY 241  
 Db 255 MDMLTFPTVEDAIKRVEMKTRTYQEVKVGMEKEEBEGNSHYVKISPLVGLDYG 314  
 QY 242 NWVPLTKEVQVYLLRLD 261  
 Db 315 DWIMPLTKEVQVYLLRLD 334  
 RESULT 2  
 CHMU YEAST  
 ID CHMU YEAST STANDARD; PRT; 256 AA.  
 AC P21178;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ChorisMat mutase (EC 5.4.99.5) (CM).  
 GN ARO7 OR OSM2 OR YPR060C OR YP9499.15C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180;  
 RX MEDLINE=89155418; PubMed=2646272;  
 RA Schmidheini T., Sperisen P., Paravicini G., Huettner R., Braus G.H.;  
 RT "A single point mutation results in a constitutively activated and  
 feedback-resistant chorisMat mutase of Saccharomyces cerevisiae.";  
 RL J. Bacteriol. 171:1245-1253 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=97313271; PubMed=9169875;  
 RA Bussey H., Storms R.K., Ahmed A., Albertmann K., Allen E., Ansoorge W.,  
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,  
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,  
 RA Chung B., Churcher C.M., Coster F., Davis K., Davis R.W.,  
 RA Dietrich F.S., Dalius H., DiPaolo T., Dubois E., Duesterhoeft A.,  
 RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,  
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,

RA Hunicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,  
 RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,  
 RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,  
 RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,  
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
 RA Urestarazu L.A., Ushinsky S., Vierendeels F., Viessers S., Voss H.,  
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.";  
 RL Nature 387:103-105 (1997).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RX MEDLINE=95062155; PubMed=7971967;  
 RA Xue Y., Lipscomb W.N., Graf R., Schnappauf G., Braus G.;  
 RT "The crystal structure of allosteric chorisMat mutase at 2.2-A  
 resolution.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10814-10818 (1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
 RX MEDLINE=96194968; PubMed=8622937;  
 RA Straeter N., Haakansson K., Schnappauf G., Braus G., Lipscomb W.N.;  
 RT "Crystal structure of the 1 state of allosteric yeast chorisMat  
 mutase and comparison with the R state.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:3330-3334 (1996).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).  
 RX MEDLINE=98046093; PubMed=9384560;  
 RA Straeter N., Schnappauf G., Braus G., Lipscomb W.N.;  
 RT "Mechanisms of catalysis and allosteric regulation of yeast  
 chorisMat mutase from crystal structures.";  
 RL Structure 5:1437-1452 (1997).  
 CC -1- CATALYTIC ACTIVITY: ChorisMat = prephenate.  
 CC a strong inhibitor. Allosterically regulated.  
 CC -1- PATHWAY: BRANCH POINT OF THE BIOSYNTHETIC PATHWAY LEADING TO THE  
 THREE AROMATIC AMINO ACIDS, PHENYLALANINE, TYROSINE, & TRYPTOPHAN.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SIMILARITY: Contains 1 chorisMat mutase domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; M24517; AAB59309.1; -;  
 DR EMBL; Z49219; CA889177.1; -;  
 DR EMBL; Z71255; CA95004.1; -;  
 DR PIR; A45921; A45921.  
 DR PDB; 1CSM; 15-SEP-95.  
 DR PDB; 2CSM; 23-DEC-96.  
 DR PDB; 3CSM; 14-JAN-98.  
 DR PDB; 4CSM; 14-JAN-98.  
 DR PDB; 5CSM; 14-JAN-98.  
 DR GeneOnline; 144325; -;  
 DR SGD; S0006264; ARO7.  
 DR GO; GO:004106; P:chorisMat mutase activity; IDA.  
 DR InterPro; IPR008238; Chor mut AroQ eu.  
 DR InterPro; IPR008951; ChorisMat mut II.  
 DR InterPro; IPR002701; ChorisMat mut.  
 DR Pfam; PF01817; ChorisMat mut; I.  
 DR PIRSF; PIRSF017318; Chor mut AroQ eu; 1.  
 KW Aromatic amino acid biosynthesis; isomerase; 3D-structure;  
 Allosteric enzyme.  
 VARIANT 226 226 T -> I (CONSTITUTIVELY ACTIVATED AND  
 FEEDBACK-RESISTANT).  
 FT TURN 3 4  
 FT HELIX 6 9  
 FT HELIX 12 34  
 FT TURN 35 35

```

FT HELIX 40 42
FT TURN 43 43
FT TURN 45 46
FT TURN 48 49
FT TURN 53 54
FT TURN 59 73
FT TURN 74 75
FT HELIX 76 78
FT TURN 80 81
FT TURN 87 89
FT TURN 106 107
FT TURN 108 110
FT HELIX 114 124
FT TURN 125 125
FT HELIX 126 129
FT TURN 137 138
FT HELIX 140 159
FT TURN 160 160
FT HELIX 161 170
FT TURN 171 171
FT HELIX 173 181
FT TURN 182 183
FT HELIX 185 191
FT TURN 192 192
FT TURN 195 211
FT TURN 212 213
FT HELIX 227 236
FT TURN 237 237
FT HELIX 238 251
FT TURN 252 252
SQ SEQUENCE 256 AA; 8C6BEBEAA3497E23 CRC64;

Query Match 37.2%; Score 498.5; DB 1; Length 256;
Best Local Similarity 40.5%; Pred. No. 7.1e-32;
Matches 102; Conservative 57; Mismatches 84; Indels 9; Gaps 4;

QY 14 VYTLASVREDLVROEDTIYGLIERAKFPSNSHTYDEKY--AQIOGCGSLVFVVKTE 71
DB 9 VLNQIRDELVRMDSIIIFKFIERSHFATCPVSVEANHPGLEIFNFKGSLDWSLDE 68
QY 72 AIOAKGRYKNPEENAFFPENLPSPVSPVSKFQFLHPGAAGININKSIWKMYFKELLP 131
DB 69 IAHSTRRESPEDETFPPDKIOKSLPINSYQILAPYAPEVNYNDKIKVYIEKIPL 128
QY 132 LA-TSGDD-GNYAQTAAANDLSLQISRIHYGKFAVVKFRDAPODYEPLTRAKDEGL 189
DB 129 ISKRGDGDKNNFGSVATRDIECLQSLRSRIHFGKFAEAKFQSDIPLYTKLIKSDVEGI 188
QY 190 MKLLTFTSVETVRKVEKKAUVFGQVNLNSDDNENRKNKDPSPVASSLYXKNWVPLTK 249
DB 189 MKNITSAVEEKLRLTKAEVIGVD-----PTNESGERRITPEVLVKIYKEIPIPTK 243
QY 250 EVQVEYLLRELD 261
DB 244 EVEVEYLLRELE 255

RESULT 3
CHMU SCHPO STANDARD; PRT; 251 AA.
AC OI3739;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable chorismate mutase (EC 5.4.99.5) (CM).
GN SPAC1658.04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Muncall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehtach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goiffau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakowski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -!- PATHWAY: BRANCH POINT OF THE BIOSYNTHETIC PATHWAY LEADING TO THE
CC -!- CATALYTIC ACTIVITY: Chorismate = prephenate.
CC -!- SUBUNIT: Homodimer (By similarity)
CC -!- SIMILARITY: Contains 1 chorismate mutase domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z98529; CAB11033.1; -.
DR PIR; T37784; T37784.
DR HSSP; P32178; 5CSM.
DR GeneDB_Spombe; SPAC1658.04c; -.
DR InterPro; IPR008238; Chor_mut_AroQ_eu.
DR InterPro; IPR008951; Chorismat_mut_II.
DR InterPro; IPR002701; Chorismat_mut.
DR Pfam; PF01817; Chorismate_mut; 1.
DR PIRSF; PIRSF017318; Chor_mut_AroQ_eu; 1.
DR Aromatic amino acid biosynthesis; Isomerase.
KW SEQUENCE 251 AA; 29050 MW; 1AC18AE4C1E6C4B7 CRC64;

Query Match 34.5%; Score 462.5; DB 1; Length 251;
Best Local Similarity 39.5%; Pred. No. 4.5e-29;
Matches 98; Conservative 54; Mismatches 85; Indels 11; Gaps 5;

QY 17 LASVREDLVROEDTIYGLIERAKFPSNSHTY---DEKYAQIOGCGSLVFVVKTEAI 73
DB 10 LENIRSAIREDTIIFNLEAQAQFPNEKYKSGKEGCLNENYDGSFNYLHHEEKV 69
QY 74 OAKAGRYKNPEENAFFPENLPSPVSPVSKFQFLHPGAAGS-ININKSIWKMYFKELLP 133
DB 70 VALVRYASPEEYP-FTDNLPELIPKFSKGFPLHPN--NVNANSEILEYVINEIVPKIS 126
QY 134 TSGDD-GNYAQTAAANDLSLQISRIHYGKFAVVKFRDAPODYEPLTRAKDEGLMKL 192
DB 127 SPGDDFDNYGTVVCGDIFCLQSLRSRIHYGKFAEAKYLANPEKYKUILARDINGINE 186
QY 193 LTFTSVETVRKVEKKAUVFGQVNLNSDDNENRKNKDPSPVASSLYXKNWVPLTKEVQ 252
DB 187 IVDAQAEERVLKELHYKALNYGRD-----AADPTKPSDRINADCVASYIKDYVIPMTKKVE 242

```

```

QY 253 VEYLLRL 260
DB 243 VYLLARL 250

RESULT 4
RIR1 HAEIN STANDARD; PRT; 756 AA.
AC P43754;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonucleoside-diphosphate reductase alpha chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
GN NRDA OR H11659.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -!- FUNCTION: Catalyzes the biosynthesis of deoxyribonucleotides from
CC the corresponding ribonucleotides, precursors that are necessary
CC for DNA synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioresoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioresoxin.
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
CC similarity).
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC large chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U32839; AAC23305.1; ALT_INIT.
CC HSSP; P00452; SRIR.
CC TIGR; H11659; -.
CC InterPro; IPR005144; ATP.
CC InterPro; IPR000788; Ribonucleo_red.
CC InterPro; IPR008926; Ribonucleo_red_N.
CC Pfam; PF03477; ATP-cone; 1.
CC Pfam; PF00317; ribonuc_red_lg; 1.
CC Pfam; PF02867; ribonuc_red_lgc; 1.
CC PRINTS; PR01183; RIBORETASEM1.
CC PROSITE; PS00089; RIBORED_LARGE; 1.
CC Oxidoreductase; DNA replication; Complete proteome.
KW ACT SITE 225 225 BY SIMILARITY.
FT ACT SITE 439 439 BY SIMILARITY.
FT ACT SITE 462 462 BY SIMILARITY.
FT SITE 749 749 INTERACTS WITH THIOREDOXIN/GLUTAREDOXIN
FT (BY SIMILARITY).

FT SITE 754 754 INTERACTS WITH THIOREDOXIN/GLUTAREDOXIN
FT (BY SIMILARITY).
SQ SEQUENCE 756 AA; 85695 MW; 839D02FA61E1B5D3 CRC64;

Query Match 7.7%; Score 103; DB 1; Length 756;
Best Local Similarity 22.9%; Pred. No. 1.9;
Matches 53; Conservative 36; Mismatches 104; Indels 38; Gaps 9;

QY 49 DEKVAIQGFCGSLVEFVVK-----NTEAIQAKAGRYKNPEENAFFPPLPSPVSPYS 102
DB 231 DLSLSINATASAIKVSQKAGINAGAIKALGSEIRGGE--AFH-----TGCIPFYK 283
QY 103 FKQF-----LHPGAASININKSIKWYFKELPLHATSGDDONTAQAANDLSLQ 153
DB 284 YFTAVKSCSQGGVGGGATLY--PIWLEAENLLVKKNGVEDNRVHRNDYGVQLNK 341
QY 154 SISRRHYGKFAEVKFRDAPQDYELPRAKQEGMLKLLTTSVEETVRKVEKAVVF 213
DB 342 LMVQRLKSGEITLFPSPDVPLGYEAFADQDK--FEELYVKYEQDPTIRKTKTVAEIF 399
QY 214 G---Q-----VNLNSDDNDNENRKFDPSVASSLYKXWVPL---TKEVQ 252
DB 400 SLLMQERASTGRYYIQNVDHCHNTHSPFDPQVAPVQSNLCLEIALPTKPLQ 450

RESULT 5
SUW3 SACDO STANDARD; PRT; 737 AA.
AC O74727;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-dependent RNA helicase SUW3, mitochondrial precursor.
GN SUW3.
OS Saccharomyces douglasii (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=46617;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CK1;
RA Golik P., Stepien P.P., Lazowska J.;
RT "The SUW3 gene of Saccharomyces douglasii is a functional equivalent
RT of its S. cerevisiae orthologue."
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Probable ATP-dependent RNA helicase involved in a
CC variety of mitochondrial post-transcriptional processes and in
CC translation. It is a key control element in nuclear-mitochondrial
CC interactions (By similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -!- SIMILARITY: Belongs to the helicase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ011586; CA009716.1; -.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR001650; Helicase_C.
CC Pfam; PF00271; helicase_C; 1.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELICC; 1.
CC Mitochondrion; Transist peptide; Helicase; ATP-binding; RNA-binding.
KW TRANSIT 1 25 MITOCHONDRION (POTENTIAL).
FT CHAIN 26 737 ATP-DEPENDENT RNA HELICASE SUW3.
FT NP BIND 239 246 ATP (POTENTIAL).
SQ SEQUENCE 737 AA; 84283 MW; E24DD4C2097CA961 CRC64;

Query Match 7.5%; Score 101; DB 1; Length 737;

```

Best Local Similarity 20.1%; Pred. No. 2.7;  
Matches 66; Conservative 48; Mismatches 93; Indels 121; Gaps 18;

QY 11 SGNVTLASVRDLVRQEDTIYGLIERAKFPNSHTYDEKYAQIQCGSLVFVVKNT 70  
DB 72 SNNVY---LRDPLFRDN-----LDRAM---QIIYNEKLSSLD-----KQVPKQL 111  
QY 71 EAIQAKGRYKQNPENAFNPENLPSPVSPKQFLHPGAAS-----ININK---SIW 121  
DB 112 AWLKLKRDSDYQOLED-----PKOOTKNYVPST--EIIYFSPGNLISLILNCKNIGNSVW 165  
QY 122 K-----MYFKELLPLLATSGD-DGNYAQTAAND--LS 150  
DB 166 KAILKNGQSNSTILDKFHVLCQTFDHMYEOLIPMTNTDDTGDGAVNDITNPAEWF 225  
QY 151 LQOSTSRRI--HYGKFAEVRKPRADQDYEPLIRAKXE-----GLMKLLFTSVSEVR 203  
DB 226 EARKIRRHIMIGP-----TNSGKTVALQKLKSVDRGYAGPLLLA-REYIDRFQ 277  
QY 204 KRVEKXAVVFGVGNLNSDDNE-----NRKPDFSVASSLY-----KXW- 243  
DB 278 SEKVRNLTGEBVIRDLDKGNPAGLISGTVEMVPINQKFDVVVLDEIQMSDADRGA 337  
QY 244 -----VIPLTKEV 251  
DB 338 WTNALLGVVSKVHLVGEKSVLPVKSI 365

RESULT 6  
RIR1\_BUCBP STANDARD; PRT; 761 AA.  
AC Q89AS4;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribonucleoside-diphosphate reductase alpha chain (EC 1.17.4.1)  
DE (Ribonucleotide reductase).  
GN NRDA OR BBP168.  
OS Buchnera aphidicola (subsp. Baizongia pistaciae).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Buchnera.  
OX NCBI\_TaxID=135842;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426901; PubMed=12522265;  
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,  
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,  
RA Tamames J., Vigueria E., Iatorre A., Valencia A., Moran F., Moya A.;  
RT "Reductive genome evolution in Buchnera aphidicola."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).  
CC -1- FUNCTION: Catalyzes the biosynthesis of deoxyribonucleotides from  
the corresponding ribonucleotides, precursors that are necessary  
for DNA synthesis (By similarity).  
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized  
thioredoxin + H(2)O = ribonucleoside diphosphate + reduced  
thioredoxin.  
CC -1- PATHWAY: DNA replication pathway; first step.  
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (By  
similarity).  
CC -1- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase  
large chain family.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AE014016; AAC26901.1; -  
CC InterPro; IPR005144; ATP.  
CC InterPro; IPR000788; Ribonucleo\_red.

InterPro; IPR008926; Ribonucleo\_red\_N.  
PFam; PF03477; ATP-cone; 1.  
PFam; PF00317; ribonuc\_red\_lg; 1.  
PFam; PF02867; ribonuc\_red\_lgc; 1.  
PRINTS; PF01183; RIBORDTASEM1.  
KW PROSITE; PS00089; RIBORED LARGE; 1.  
Oxidoreductase; DNA replication; Complete proteome.  
FT ACT\_SITE 225 225 BY SIMILARITY.  
FT ACT\_SITE 439 439 BY SIMILARITY.  
FT ACT\_SITE 462 462 BY SIMILARITY.  
FT ACT\_SITE 754 754 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN  
(BY SIMILARITY).  
FT SITE 759 759 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN  
(BY SIMILARITY).  
FT SITE 759 759 INTERACTS WITH THIOREDOXIN/ GLUTAREDOXIN  
(BY SIMILARITY).  
SQ SEQUENCE 761 AA; 87124 MW; 7FB14E136283BDCF CRC64;

Query Match 7.1%; Score 94.5; DB 1; Length 761;  
Best Local Similarity 20.8%; Pred. No. 9;  
Matches 46; Conservative 42; Mismatches 104; Indels 29; Gaps 8;

QY 49 DEKYAQIQCGSLVFVVKNTAIOAKGRYK---NPEENAFNPENLPSPVSPKQ 105  
DB 231 EDNLNSINATTSATVKVYSQRA-GIGINAGRIRALGSPIRNG---DTLHTGCIFFYKHQ 286  
QY 106 F-----LHPCAASININKSIWYFKELLPLLATSGDGNYAQTAANDLSLQSI 156  
DB 287 SAVKSCSQGGVGGAAATIFY--PIWHFEIESLLVLRKNGRIENVRVHMYSVQLNKLNY 344  
QY 157 RRIHYGKFAEVRKPRADQDYEPLIRAKDEGLMKLLTFTSVETVRKRVKKAWFG-- 214  
DB 345 QRLILGEHITLFPSPDVPNLYDSFPNNQEK--PERLYKYENDKTKIRKKVKASYLFSLM 402  
QY 215 -QE-----VNLNSDDNENRKPDPSPVASSLYKNWVPLT 248  
DB 403 MQERTSTGTIRYIVNVHCHNSHSAFNPKIPIROSNLCLEIT 443

RESULT 7  
YFLL\_YEAST STANDARD; PRT; 295 AA.  
ID YFLL\_YEAST  
AC P43613;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical 34.2 kDa protein in SAPI55-YMR31 intergenic region  
DE precursor.  
GN YFR041C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288c / AB972;  
RX MEDLINE=95400292; PubMed=7670463;  
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,  
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,  
RA Yamazaki M., Tashiro H., Eki T.;  
RT "Analysis of the nucleotide sequence of chromosome VI from  
Saccharomyces cerevisiae."  
RL Nat. Genet. 10:261-268(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=S288c / AB972;  
RX MEDLINE=96287652; PubMed=866379;  
RA Eki T., Naitou M., Hagiwara H., Ozawa M., Sasanuma S.-I.,  
RA Sasanuma M., Tsuchiya Y., Shibata T., Hanaoka F., Murakami Y.;  
RT "Analysis of a 36.2 kb DNA sequence including the right telomere of  
chromosome VI from Saccharomyces cerevisiae."  
RL Yeast 12:149-167(1996).  
CC -1- SIMILARITY: Contains 1 J domain.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



```

RN SEQUENCE FROM N.A.
RP Wedler H., Scharfe M., Wedler E., Wambutt R.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN
RP SEQUENCE OF 1-19 FROM N.A.
RA Hernandez K., Weber N., Wipfli P., Schmidheini T.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN
RP SUBCELLULAR LOCATION, AND FUNCTION.
EX MEDLINE=22016919; PubMed=12022229;
RA Oeffinger M., Lueng A., Lamond A., Tollervoy D.;
RT "Yeast Pescadillo is required for multiple activities during 60S
RL ribosomal subunit synthesis.";
RL RNA 8:626-636(2002).
CC -!- FUNCTION: May function both in ribosome synthesis and in cell-
CC cycle regulation.
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar.
CC -!- SIMILARITY: Contains 1 BRCT domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 272888; CAA97106.1; -.
CC PIR; S64410; S64410.
CC GeroOnline; 141415; -.
CC SGD; S0003335; RRP13.
CC GO; GO:0005730; C:nucleolus; IDA.
CC GO; GO:0042273; P:ribosomal large subunit biogenesis; IDA.
CC InterPro; IPR001357; BRCT.
CC Pfam; PF00533; BRCT; 1.
CC SMART; SM00292; BRCT; 1.
CC RIBOSOME biogenesis; Nuclear protein; Coiled coil.
CC PROSITE; PS50172; BRCT; 1.
CC KW Ribosome biogenesis; Nuclear COILED COIL (POTENTIAL).
CC FT DOMAIN 291 346
CC FT DOMAIN 355 449
CC FT DOMAIN 526 604
CC FT DOMAIN 560 580
CC FT DOMAIN 479 496
CC FT DOMAIN 501 515
CC FT DOMAIN 501 515
CC SQ SEQUENCE 605 AA; 69877 MW; 131001C956787BE5 CRC64;
CC
CC Query Match
CC Best Local Similarity 6.9%; Score 92; DB 1; Length 605;
CC Matches 62; Conservative 50; Mismatches 116; Indels 74; Gaps 13;
CC
CC 2 AKAAEQSPGNGVYTLASVRE-----DLVRQSDTIYGLIERAKPPNSHTYDEKYAQ 54
CC 102 AKRLEENRDS---YTLDHIIKERYPSPFPDAIRDIDALNMLFLFSLPSTNQVSSKIIND 158
CC
CC 55 IQGFCGSLVFVVKV-----TEAIQAKAGRYKKNPEAFPEENLPESI-- 97
CC 159 AQKICNQLAVAKERLVKVFVSIKGVYQANIKGEEVRLVFPK---FPENISDVDF 215
CC
CC 98 VPSYFVKQFLHPGAASININKSIWYFKLLPLLATSGDGNAYQATANDLSLQSTSR 157
CC
CC 216 RIMLTFLEFY-----STLLHFVLYKLY-----TDSGLIYPPKLDLKKDKIIS- 257
CC
CC 158 RIHYGKFAEVKVPDPAQDVEPL-----IRAKKEGLMKLTLF-----TSVEETR 203
CC
CC 258 --GLSSVILSRQEDSLKLDPTIEIEDVKVSLDASTLSALNADANTDETEKEEGR 315
CC
CC 204 KRVEK-KAVFGQEVNLSDDNDNEN-----KFDPSVASSLYKNWVPLTKEVQVE 254
CC
CC 316 KKQKEQEQEKNQETELDTEDNNKNGDILLIQPSKYDSPVA-SLPSAFVYVVSREPID 374
CC
CC 255 YL 256

```

```

Db 375 IL 376
RESULT 10
LON BRUME
ID LON BRUME STANDARD; PRT; 812 AA.
AC Q8YHC6;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ATP-dependent protease La (EC 3.4.21.53).
GN LON OR BMEI0876.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=23459;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=16W / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapural V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -!- FUNCTION: Degrades short-lived regulatory and abnormal proteins in
CC presence of ATP. Hydrolyzes two ATPs for each peptide bond cleaved
CC in the protein substrate (By similarity).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of large proteins such as globin,
CC casein and denaturated serum albumin, in presence of ATP.
CC -!- SUBUNIT: Homotetramer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family S16.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009528; AAL52057.1; ALT_INIT.
CC PIR; AF3361; AF3361.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR003593; AAA_ATPase_cent.
CC InterPro; IPR001270; Chaprinin_c1pA/S.
CC InterPro; IPR001984; Peptidase_S16.
CC InterPro; IPR008268; Peptid_S16_AS.
CC InterPro; IPR008269; Peptid_S16_C.
CC InterPro; IPR004815; Pept_S16_Lon.
CC InterPro; IPR003111; Pept_S16_N.
CC Pfam; PF00004; AAA; 1.
CC Pfam; PF02190; LON; 1.
CC Pfam; PF05362; Lon_C; 1.
CC PRINTS; PR00300; CLIPPROTEASEA.
CC PRINTS; PR00830; ENDOLAPTAASE.
CC SMART; SM00382; AAA; 1.
CC SMART; SM00464; LON; 1.
CC TIGRfams; TIGR00763; Lon; 1.
CC PROSITE; PS01046; LON_SER; 1.
KW Hydrolase; Serine protease; ATP-binding; Complete proteome.
FT NP_BIND 367 374 ATP (POTENTIAL).
FT ACT_SITE 689 689 BY SIMILARITY.
SQ SEQUENCE 812 AA; 89895 MW; 401E15F1FDD24E98 CRC64;
CC
CC Query Match
CC Best Local Similarity 6.9%; Score 92; DB 1; Length 812;
CC Matches 63; Conservative 44; Mismatches 85; Indels 94; Gaps 14;

```



QY 4 AAEQSPDSGNVY---TLASVREDLVQEDTIYGLIE---RAKFP--SNSHTYDEKIAQI 55  
 DB 69 AADDDPAPDAIYEIGTIANVLQ--LLKLPDGTVKLVVEGTARAKISKFTDREDYHEAYA-- 125  
 QY 56 QGFCGSLVFEVVKNTBAIQAKAGRYKNPEENAFFPENLPPSVPSYKQFLHPCGAASIN 115  
 DB 126 -----AALQ-----EPEDAVEIEALARSVVS--DFENY-----VK 154  
 QY 116 INKSIWMTFKELLPLLATSGDGNVQAOTANDLSLOSISRRIHYGKFAEVKFRDAPQ 175  
 DB 155 LNKKI-----SPEVWGAAQSIDDYSLADTVASHLAI-----KIPE 190  
 QY 176 DYEPIIRAKDKGLMKLLTF-----TSVETVKRVEKKAIVFGQVNLN----- 220  
 DB 191 KOEMLSVLSVRERLEKALSFMEAEISVLQVKRISRVKQMKVQREYLYLNEQMKAIQK 250  
 QY 221 -----SDNDNENRKFPDPSVASSLYKNWVPLTKVQVEYL--LRRL 260  
 DB 251 ELGDSGDEGRDEVAIEERITKT-----KLSKEAREKALAEKCL 289

RESULT 11  
 LON\_BRUSU  
 ID LON\_BRUSU STANDARD; PRT; 812 AA.  
 AC QSG017;  
 DT 15-MAR-2004 (Rel. 43, Created)  
 DT 15-MAR-2004 (Rel. 43, Last sequence update)  
 DE ATP-dependent protease La (EC 3.4.21.53).  
 GN LON OR BR1106.  
 OS Brucella suis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29461;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1330 / Biovar 1;  
 RX MEDLINE=22247741; PubMed=12271122;  
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
 RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,  
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,  
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,  
 RA Hoover D.B., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;  
 RT "The Brucella suis genome reveals fundamental similarities between  
 animal and plant pathogens and symbionts."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).  
 CC -!- FUNCTION: Degrades short-lived regulatory and abnormal proteins in  
 presence of ATP. Hydrolyzes two ATPs for each peptide bond cleaved  
 in the protein substrate (By similarity).  
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of large proteins such as globin,  
 casein and denaturated serum albumin, in presence of ATP.  
 CC -!- SUBUNIT: Homotetramer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to peptidase family S16.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC ENBL; AE014412; AAN30026.1; -  
 DR TIGR; BR1106; -  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003593; AAA\_ATPase\_centre.  
 DR InterPro; IPR001270; Chaprinin\_c1pA/B.  
 DR InterPro; IPR001984; Peptidase\_S16.  
 DR InterPro; IPR008268; Peptidase\_S16.  
 DR InterPro; IPR008268; Peptidase\_S16.  
 DR InterPro; IPR008268; Peptidase\_S16.  
 DR InterPro; IPR004815; Pept\_S16\_Lon.

InterPro; IPR003111; Pept\_S16\_N.  
 Pfam; PF00004; AAA; 1.  
 Pfam; PF02190; LON; 1.  
 Pfam; PF05362; Lon\_C; 1.  
 PRINTS; PRO0300; CLPPROTEASEA.  
 PRINTS; PRO0830; ENDOLAPTASE.  
 SMART; SM00382; AAA; 1.  
 SMART; SM00464; LON; 1.  
 TIGRFAMs; TIGR00763; Lon; 1.  
 PROSITE; PS01046; LON\_SER; 1.  
 Hydrolyase; Serine protease; ATP-binding; Complete proteome.  
 NP\_BIND 367 374 ATP (POTENTIAL).  
 FT ACT\_SITE 689 689 BY SIMILARITY.  
 SQ SEQUENCE 812 AA; 89794 MW; 0FOAC318C9227FBF CRC64;  
 Query Match 6.9%; Score 92; DB 1; Length 812;  
 Best Local Similarity 22.0%; Pred. No. 15;  
 Matches 63; Conservative 44; Mismatches 85; Indels 94; Gaps 14;  
 QY 4 AAEQSPDSGNVY---TLASVREDLVQEDTIYGLIE---RAKFP--SNSHTYDEKIAQI 55  
 DB 69 AADDDPAPDAIYEIGTIANVLQ--LLKLPDGTVKLVVEGTARAKISKFTDREDYHEAYA-- 125  
 QY 56 QGFCGSLVFEVVKNTBAIQAKAGRYKNPEENAFFPENLPPSVPSYKQFLHPCGAASIN 115  
 DB 126 -----AALQ-----EPEDAVEIEALARSVVS--DFENY-----VK 154  
 QY 116 INKSIWMTFKELLPLLATSGDGNVQAOTANDLSLOSISRRIHYGKFAEVKFRDAPQ 175  
 DB 155 LNKKI-----SPEVWGAAQSIDDYSLADTVASHLAI-----KIPE 190  
 QY 176 DYEPIIRAKDKGLMKLLTF-----TSVETVKRVEKKAIVFGQVNLN----- 220  
 DB 191 KOEMLSVLSVRERLEKALSFMEAEISVLQVKRISRVKQMKVQREYLYLNEQMKAIQK 250  
 QY 221 -----SDNDNENRKFPDPSVASSLYKNWVPLTKVQVEYL--LRRL 260  
 DB 251 ELGDSGDEGRDEVAIEERITKT-----KLSKEAREKALAEKCL 289

RESULT 12  
 ATPA\_BUCAP  
 ID ATPA\_BUCAP STANDARD; PRT; 510 AA.  
 AC O51874;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ATP synthase alpha chain (EC 3.6.3.14).  
 GN ATPA OR BUSG006.  
 OS Buchnera aphidicola (subsp. Schizaphis graminum).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Buchnera.  
 OX NCBI\_TaxID=98794;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97361981; PubMed=9216881;  
 RA Clark M.A., Baumann P.;  
 RT "The (F1F0) ATP synthase of Buchnera aphidicola (endosymbiont of  
 aphids): genetic analysis of the putative ATP operon."  
 RL Curr. Microbiol. 35:84-89(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98184963; PubMed=9516544;  
 RA Clark M.A., Baumann L., Baumann P.;  
 RT "Sequence analysis of a 34.7-kb DNA segment from the genome of  
 Buchnera aphidicola (endosymbiont of aphids) containing groEL, dnaA,  
 the ATP operon, gldA, and rho."  
 RL Curr. Microbiol. 36:158-163(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22084549; PubMed=12089438;  
 RA Tamás I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,  
 Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;





ALP6 SCHPO  
ID ALP6 SCHPO STANDARD; PRT; 832 AA.  
AC Q9USQ2; Q94366; Q9P954;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Spindle pole body component alp6 (Altered polarity protein 6).  
GN ALP6 OR SPBC428.20C OR SPBC902.01C.  
OS Schizosaccharomyces pombe (Fission Yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RP MEDLINE=20532503; PubMed=11080156;  
RA Vardy L., Toda T.;  
RT "The fission yeast gamma-tubulin complex is required in G(1) phase and  
RT is a component of the spindle assembly checkpoint.";  
RL EMBO J. 19:6098-6111(2000).  
RN [2]  
SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
RA Welter J., Vanstreels E., Rieger M., Schaefer M., Moestl D., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fricz C., Holzer E., Wambutt R., Purnelle B.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Reinhardt K., Pohl T.M.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Huxst S.M.,  
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe.";  
RL Nature 415:871-880(2002).  
CC -!- FUNCTION: Component of the gamma tubule complex that is required  
CC for the regulation of both interphase microtubules and mitotic  
CC bipolar spindles.  
CC -!- SUBCELLULAR LOCATION: Spindle pole body and the microtubule  
CC organizing center (MTOC).  
CC -!- SIMILARITY: Belongs to the GCP family.  
CC -!- CAUTION: Ref.1 sequence differs from that shown due to a  
CC frameshift in position 1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; AB040811; BA994097.1; ALT\_FRAME.  
DR ENBL; AL333066; CAB2095.1; -  
DR ENBL; AL034382; CA22295.1; -  
DR GeneDB SPombe; SPBC428.20C; -  
DR InterPro; IPR007259; SPC97\_Spc98.  
DR Pfam; PF04130; SPC97\_Spc98; 1.

KW Microtubule; Mitosis.  
FT CONFLICT 665 675 MISSING (IN REF. 2).  
SQ SEQUENCE 832 AA; 95996 MW; C92771C3DBF5COLA CRC64;  
  
Query Match 6.8%; Score 91.5; DB 1; Length 832;  
Best Local Similarity 25.4%; Pred. No. 17;  
Matches 45; Conservative 28; Mismatches 67; Gaps 11;  
  
QY 108 HPGNAININKSIWK-MYF---KELLPLATS-----GDDGVYAQTAAANDLSLQSI 155  
Db 408 HQGGQDV-----VWKGKYLFDKELIFSLSEELVDKIFLGKSLNFARYGCGDFDWAQE- 461  
QY 156 SRRHYGKFAEVKFRDAPDQYELPRARAKDEGLMKLLTFTSVEE---TVKRVEKAV 211  
Db 462 ----HYQKLVKLSYRD-FHSLETVDVKAYTESINHLVYL--MEEVFLHTDHLKAKKYL 514  
QY 212 VFQGE--VNLSDDNDNE-----NRKPDPSVASSLYKNWVPLTKEVQVEYLLRLD 261  
Db 515 LLGQGGDFVDLLMESLGNLSLDQPANTLFRHNLTAAL--ESAIRSSNASEYEYVLRKLD 570  
  
RESULT 15  
HSLU CAMJE STANDARD; PRT; 439 AA.  
ID HSLU CAMJE  
AC Q9PHLO;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE ATP-dependent hsl protease ATP-binding subunit hslu.  
GN HSLU OR CJO662C.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
OC Campylobacteraceae; Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=NTCT 11168;  
RX MEDLINE=20150912; PubMed=10689204;  
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
CC -!- FUNCTION: Chaperone subunit of a proteasome-like degradation  
CC complex (By similarity).  
CC -!- SUBUNIT: A double ring-shaped homohexameric (By similarity).  
CC each side by a ring-shaped hslu homohexamer (By similarity).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: Belongs to the clpX chaperone family. Hslu subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC ENBL; AL139075; CAB75294.1; -  
DR PIR; H81414; H81414.  
DR HSP; P32168; 1D02.  
DR HAMAP; MF 00249; - 1.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR003959; AAA\_ATPase\_Centr.  
DR Pfam; PF00004; AAA; 1.  
DR SMART; SM00382; AAA; 1.  
DR TIGRFAMs; TIGR00390; hslu; 1.  
KW Chaperone; ATP-binding; Complete proteome.  
FT NP\_BIND 56 63 ATP (POTENTIAL).

SQ SEQUENCE 439 AA; 49779 MW; 36FE3914DFBE48DA CRC64;  
Query Match 6.8%; Score 91; DB 1; Length 439;  
Best Local Similarity 25.2%; Pred. No. 8.5;  
Matches 53; Conservative 31; Mismatches 76; Indels 50; Gaps 11;  
QY 51 KYAQIQGFCGSLVEFVVKNTAIOAKAG-----RYKNPEENAFFPENLPPSIVPSYS 102  
Db ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||  
84 KYTEV-GFVGRDVESWVRL---ANAALNLVKNQREKKNKDKIDEFIEN-----KI 130  
QY 103 FKQFLHFGAASININKSIWKNYFKELLPLLAT---SGD-----DGNVAQTAAN-- 147  
Db ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||  
131 LEKLLPPLPKGISDEK---QEEYKNSLEKQRTKLRNGDLDESTIEIISQNWFDINPILP 187  
QY 148 -DLSLIQISRRRIHYG--KFVAEVKFRDAPQDYELIRAKDXEGLMKLLTFTSVETVRK 204  
Db ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||  
188 PENGAMQDIVKIVGVGSKVKVKKEMKIDAK-----NALKNEAGEKILDQESIKSEALK 240  
QY 205 RVEKKAVVFGQEVNLSDDNDNENKFDPS 234  
Db ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||  
241 RAENEGIIFIDEIDKIAVSSGNSNRQ-DPS 269

Search completed: September 25, 2004, 02:53:01  
Job time : 20 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:55:17 ; Search time 4193 Seconds

(without alignments)  
2697.957 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDSGNVYTLASV.....NWVPLTKVEQVEYLLRLD 261

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US10624061/runat\_24092004\_092406\_16911/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptO -NORM=ext -HAPSPZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10624061@cgn2\_1.1.3731 @runat\_24092004\_092406\_16911 -NCPUs=6 -ICPU=3  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.rc:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.or:\*  
21: em.ov:\*  
22: em.pat:\*  
23: em.ph:\*  
24: em.pl:\*  
25: em.ro:\*  
26: em.scs:\*  
27: em.un:\*  
28:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1339	100.0	1020	6	AR404688 Sequence
2	818	61.1	798	6	AX506277 Sequence
3	818	61.1	829	8	AY133840 Arabidops
4	818	61.1	993	6	AX343935 Sequence
5	818	61.1	1006	6	AR236634 Sequence
C 6	818	61.1	1006	6	AR236640 Sequence
7	818	61.1	1006	6	AX343933 Sequence
8	818	61.1	1006	6	AX464575 Sequence
C 9	818	61.1	1006	6	AX464581 Sequence
10	818	61.1	1006	6	AY052238 Arabidops
11	818	61.1	1055	8	AY052238 Arabidops
12	767.5	57.3	525	6	AR404683 Sequence
13	698	52.1	1005	6	AX506042 Sequence
14	696	52.0	1207	6	AR236633 Sequence
C 15	696	52.0	1207	6	AR236639 Sequence
16	696	52.0	1207	6	AX464574 Sequence
C 17	696	52.0	1207	6	AX464580 Sequence
18	696	52.0	1207	8	ATCHMT
19	691	51.6	1000	8	L47356
20	686	51.2	1348	8	AK105512 Oryza sat
21	685.5	51.2	1231	6	AR404687 Sequence
22	685.5	51.2	1250	8	AK068983 Oryza sat
23	678.5	50.7	1435	8	AK101220 Oryza sat
24	677.5	50.6	1274	8	AY089156 Arabidops
25	670.5	50.1	1223	6	AR404686 Sequence
26	669.5	50.0	951	8	BT005306 Arabidops
27	669.5	50.0	1143	8	AK117860 Arabidops
28	654.5	48.9	1217	6	AR236635 Sequence
C 29	654.5	48.9	1217	6	AR236641 Sequence
30	654.5	48.9	1217	6	AX464576 Sequence
C 31	654.5	48.9	1217	6	AX464582 Sequence
32	654.5	48.9	1217	8	AF131219 Arabidops
33	646	48.2	1043	8	AK069725 Oryza sat
34	615	45.9	81062	2	AC128660 Medicago
35	602	45.0	780	6	AR404689 Sequence
36	602	45.0	780	8	BT009499 Triticum
37	587	43.8	5176	6	AR236637 Sequence
38	587	43.8	5176	6	AX464578 Sequence
39	587	43.8	5177	8	AT242648 Arabidops
C 40	573.5	42.8	98473	8	AT130N20 Arabidops
41	549.5	41.0	5099	6	AR236636 Sequence
42	549.5	41.0	5099	6	AX464577 Sequence
43	549.5	41.0	5099	8	ATH242647 Arabidops
C 44	549.5	41.0	77401	8	AB026657 Arabidops
C 45	540	40.3	93802	8	AC018364 Arabidops

# ALIGNMENTS

RESULT 1

AR404688  
 LOCUS AR404688 1020 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 15 from patent US 6627798.  
 ACCESSION AR404688  
 VERSION AR404688.1 GI:40153356  
 KEYWORDS Unknown.  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1020)  
 AUTHORS Falco, S.C., Fancdu, O.O. and Lee, J.-M.  
 TITLE Aromatic amino acid biosynthetic enzymes  
 JOURNAL Patent: US 6627798-A 15 30-SEP-2003;  
 FEATURES  
 source  
 1. 1020  
 Location/Qualifiers  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,57e-121 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-624-061-16 (1-261) x AR404688 (1-1020)  
 QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrIleuAlaSerVal 20  
 Db 23 ATGGCCAAAGCAGCAGAAACAAGTCTGATTCGGAAATGTGTACACGCTAGCTTCTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
 Db 83 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGTCTCTCATTCAGAGAGCCNAG 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 Db 143 TTCCTTAGCAATTCACACCTATGATGAAGATGCTCAATCCAGGGTTTTGTGGC 202  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 Db 203 TCATTGGTGGATTTGTTTAAAGATACAGAGCCATTCAGCTAAGCTAAGCTGGAAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPheProGluAsnLeuProProSerIleValProSer 100  
 Db 263 AAAAACCCCTGAAGAAACGCCCTTCTCCAGAAAATTTACCACCATCAATTGTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 Db 323 TACTCCTTCAACAGAGTTTTCATCCCTGGTCTGCTCTCAATTAACAATAAACAAGTCCATC 382  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuAlaThrSerGlyAspAspGlyAsn 140  
 Db 383 TGGAAAATGTATTTCAAGAGTTACTTCCATTTGCTTCTACTTCTGGGTGATGATGCCAAC 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgAlaIleHis 160  
 Db 443 TATGGCAACTCGAGCTAATGACCTTTCATTTATGTCAGTCCATCTCTAGAAGGATTCAC 502  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 Db 503 TATGGAAAGTTGTAGCTGAGTGAATTCAGGGATCTCTCAGACTACGAGCCCTTTA 562  
 QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
 Db 563 ATTCGAGCTAAGGATAAAGAGGATTCATGAATTTGTGACATTTAACAAGCGTTGAAGAG 622  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 Db 623 ACGGTGAGGAAGAGAGTTGAAGAGAGGCTGTGCTGTTTGGCAGGAGAGTGAATCTTAAC 682

QY 221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaLaserSerLeuTyr 240  
 Db 683 AGTGATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGACATGAC 742  
 QY 241 LysAsnTrpValIlePheProLeuThrLysGluValGlnValGlnValGlnValGlnValGln 260  
 Db 743 AAAAATTGGGTGATACCTCTCACCAAGAGGTTTCAAGTTGAGTACCTCTTTCGCCGCTCTA 802  
 QY 261 Asp 261  
 Db 803 GAC 805  
 RESULT 2  
 AX506277  
 LOCUS AX506277 798 bp DNA linear PAT 27-SEP-2002  
 DEFINITION Sequence 972 from Patent WO0216655.  
 ACCESSION AX506277  
 VERSION AX506277.1 GI:23387514  
 KEYWORDS  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 REFERENCE 1  
 AUTHORS Harper, J.F., Krops, J., Wang, X. and Zhu, T.  
 TITLE Stress-regulated genes of plants, transgenic plants containing  
 same, and methods of use  
 JOURNAL Patent: WO 0216655-A 972 28-FEB-2002;  
 The Scripps Research Institute (US); Syngenta Participations AG  
 (CH)  
 FEATURES  
 Location/Qualifiers  
 1. 798  
 /organism="Arabidopsis thaliana"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:3702"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,29e-70 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservatives: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 6 Gaps: 2  
 US-10-624-061-16 (1-261) x AX506277 (1-798)  
 QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrIleuAla 18  
 Db 1 ATGGCAAGAGTCTTCCGAATCGGATTCGGGTCTCTGTTTCCCAATGTACTCAGCTTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
 Db 61 TTAATCAGAGATTCCTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTTCATCGAGAGA 120  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 121 GCTAAGTTTCCACTCAATTCCTGCTTTCGAGGAATCTGTTGCTAGATTCCTGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 181 TTCTCTTCTCCTAGTATTTTCGTGAGAGACAGAAATCATCCAAAGCTAAGGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98  
 Db 241 AGATATGAATACCCGGAAGAGAAATCCCTTCTTCCTTCGAGAACATTCCTCACTCGGTTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 301 CCTACGACAAATATCCATCGGCTTTCACCCCAAGCTCTATCTGTAAACATTAACAAA 360  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138







SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,  
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.  
TITLE Methods and compositions for the modulation of chorismate synthase  
and chorismate mutase expression or activity in plants  
JOURNAL Patent: US 6465217-A 3 15-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 1..73e-70 Length: 1006  
Score: 818.00 Matches: 162  
Percent Similarity: 72.83% Conservative: 31  
Best Local Similarity: 61.13% Mismatches: 68  
Query Match: 61.09% Indels: 4  
DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x AR236634 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
DB 64 ATGGCAAGAGCTTCGAATCGGATTCGGGTTCGTGTTTCCAAATGATCTGAGTCTTGAC 123  
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
DB 124 TTAATCAGAGATCGTTGATTAGCGAAGACACCATCGTCTTCAGTTGATCGAGAGA 183  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
DB 184 GCTAAGTTTCCACTCAATCTCTGCTTTCGAGGAATCTCGTTGTCTAGATCTGGAAGT 243  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
DB 244 TTCTCTTCTCTACTGAGTTTTTCGTCAGAGACAGAAATCATCCAACTAGGTAGGA 303  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
DB 304 AGATATGAATACCCGGAAGAGATCCTTCTCTTCCCTTGAGAAATCTCTCACTCGGTTTTT 363  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
DB 364 CCTACGCACAAATATCCATCGGCTTTTGCACCCCTAAGGCTCTATCTGTTAACATTAA 423  
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
DB 424 CAATCTGGGATATTACTTTAAGAAATGCTTCTTCTTGTTCACAACTGGCGATGAT 483  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158  
DB 484 GGCAACTATCCATCACTGCTGATGATCTCGCTGTTTACAGCTCTCTTCGAGAAGG 543  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
DB 544 ATTCACTACGGTAAATTTGAGTGTAGTCAATTCAGAGATGCTCCACAGATACAGAG 603  
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
DB 604 CTTGCGATTCCGCTCAGGATAGAGAGCTTTGATGAAGCTGTGACGTTTGAGAAAGTA 663  
QY 199 GluGluThrValArgLysArgValGluLysValAlaValPheGlyGlnGluValAsn 218  
DB 664 GAGAAATGTTTAAAGAGAGGTGCGAAGAAAGCAGAAACGTTTGGACAAAGATGAAA 723  
QY 219 LeuAsnSerAspAsnAsnGluAsnArgLysPhe-----AspProSerValAla 236  
DB 724 TTCAACTCTGCTATGCGGATGAGAGTAAAGAGAGATATAAAGTGGATCCATTGCTTGC 783

QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
DB 784 TCTCGCATCTACGGGAATGCTTATCCCTCTCACTAAGCTCGTGTAGGTGAGTATCTT 843  
QY 257 LeuArgArgLeuAsp 261  
DB 844 CTACGTGCTCTCGAT 858  
RESULT 6  
AR236640/c  
LOCUS AR236640 1006 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 9 from patent US 6465217.  
ACCESSION AR236640  
VERSION AR236640.1 GI:27280726  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1006)  
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,  
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.  
TITLE Methods and compositions for the modulation of chorismate synthase  
and chorismate mutase expression or activity in plants  
JOURNAL Patent: US 6465217-A 9 15-OCT-2002;  
FEATURES Location/Qualifiers  
source 1..1006  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 1..73e-70 Length: 1006  
Score: 818.00 Matches: 162  
Percent Similarity: 72.83% Conservative: 31  
Best Local Similarity: 61.13% Mismatches: 68  
Query Match: 61.09% Indels: 4  
DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x AR236640 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
DB 943 ATGGCAAGAGCTTCGAATCGGATTCGGGTTCGTGTTTCCAAATGATCTGAGTCTTGAC 884  
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
DB 893 TTAATCAGAGATCGTTGATTAGCGAAGACACCATCGTCTTCAGTTGATCGAGAGA 824  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
DB 823 GCTAAGTTTCCACTCAATCTCTGCTTTCGAGGAATCTCGTTGTCTAGATCTTGGAGT 764  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
DB 763 TTCTCTTCTCTACTGAGTTTTTCGTCAGAGACAGAAATCATCCAACTAGGTAGGA 704  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
DB 703 AGATATGAATACCCGGAAGAGATCCTTCTCTTCCCTTGAGAAATCTCTCACTCGGTTTTT 644  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
DB 643 CCTACGCACAAATATCCATCGGCTTTGACCCCTAAGGCTCTATCTGTTAACATTAA 584  
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
DB 593 CAATCTGGGATATTACTTTAAGAAATGCTTCTTCTTGTTCACAACTGGCGATGAT 524  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158  
DB 523 GGCAACTATCCATCACTGCTGATGATCTCGCTGTTTACAGCTCTTTCGAGAAGG 464  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178

```

Db      463 ATTCACTACGGTAAATTTGTAGCTAGGTCAAAATTCAGAGATGCTCCCAAGATTACGAG 404
QY      179 ProLeuileArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db      403 CTGCGGATTCGGCTCAGATAGAGAGGCTTTGATGAGAGCTGTGACCGTTTGAGAAAGTA 344
QY      199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218
Db      343 GAAGAAATGTTTAAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTGGCAAGAAAGTAAA 284
QY      219 LeuAsnSerAspAspAsnAspGluAsnArgLysPhe-----AspProSerValAla 236
Db      283 TTCAACTCTGGCTATGGCGATGAGATGAGAGAGATGATATAAGTGCATCCATTGCTTGC 224
QY      237 SerSerLeuTyrLysAsnTrpValleProLeuThrLysGluValGlnValGluTyrLeu 256
Db      223 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 164
QY      257 LeuArgArgLeuAsp 261
Db      163 CTACGTCGTCGAT 149

RESULT 7
AX343933      1006 bp      DNA      linear      PAT 01-FEB-2002
LOCUS
DEFINITION
Sequence 3 from Patent WO0200901.
ACCESSION
AX343933
VERSION
AX343933.1 GI:18491972
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
Badur, R., Geiger, M., Kunze, I. and Sommer, S.
Changing the fine chemical content in organisms by genetically
modifying the shikimate pathway
Patent: WO 0200901-A 3 03-JAN-2002;
Sungene GmbH & Co. KGaA (DE)
FEATURES
source
1..1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
64..861
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD22398.1"
/db_xref="GI:18491973"
/db_xref="REMBL:CAD22398"
/translation="MARVFESDGGSCNVLSLDLRESLIHQEDTIVFSLIERAKPP
LNSPAPESRLDGSFSLTEFFVRETEIIQAKGVYEPENPFLENIPHSVFT
HKYPSALPKALSVINQKINDIVFKEILPLFKPGDDGNVPSSTAADLACLQALSR
IHYGFVEVFRDAPQYEPALRAQDREALKLLTPKEVEYVKKVQKAEFFQGE
VAFNSGYGDESKKTKVDPLLASRIYGEWLPLTKLVEVEYLLRLD"

ORIGIN
Alignment Scores:
Pred. No.:      1..73e-70      Length:      1006
Score:          818.00      Matches:      162
Percent Similarity: 72.8%      Conservative: 21
Best Local Similarity: 61.1%      Mismatches: 68
Query Match:      61.09%      Indels:      4
DB:              6      Gaps:      2

US-10-624-061-16 (1-261) x AX343933 (1-1006)

QY      1 MetaLysAlaAlaGluLysSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db      64 ATGGCAAGAGATTTCGAATTCGAGTTCGGGTTCTGCTGTTCCAAATGACTAGTCTTTC 123

```

```

QY      19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38
Db      124 TTAATCAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTCAGCTGTGATCGAGAGA 183
QY      39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58
Db      184 GCTAAGTTTCCACTCAATTCCTCTGCTTCGAGGAATCTCGTTGCTAGATTCTGGAAGT 243
QY      59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db      244 TTCTCTCTCTCACTAGTTTTTCGTCAGAGAGACAGAAATCATCCAAGCTAAGTAGGA 303
QY      79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db      304 AGATATGATACCCGGAAGAGATCCCTTCTCTCTGAGAAATCTCTCACTCGGTTTTT 363
QY      99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118
Db      364 CCTACGCAAAATATCCATCGGCTTTGCACCCCTAAGGCTCTATCTGTTAACATTACCAA 423
QY      119 SerIleTyrPheLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138
Db      424 CAAATCTGGGATATTTACTTTAAAGAAATTCCTTTGTTGTCAAACCTGGCGATGAT 483
QY      139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158
Db      484 GGCAACTATCCATCACTGCTGCTAGTATCTCGCTGTTTACAAGCTCTTTCGAGAAAG 543
QY      159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178
Db      544 ATTCACTACGGTAAATTTCTAGCTAGGTCAAATTCAGAGATGCTCCCAAGATTACGAG 603
QY      179 ProLeuileArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db      604 CTGCGGATTCGGCTCAGATAGAGAGGCTTTGATGAAGCTCTTGACGTTTGAGAAGTA 663
QY      199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218
Db      664 GAAGAAATGTTTAAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTGGCAAGAAAGTAAA 723
QY      219 LeuAsnSerAspAspAsnAspGluAsnArgLysPhe-----AspProSerValAla 236
Db      724 TTCAACTCTGGCTATGGCGATGAGATGAGAGAGATATAAGTGCATCCATTGCTTGC 783
QY      237 SerSerLeuTyrLysAsnTrpValleProLeuThrLysGluValGlnValGluTyrLeu 256
Db      784 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843
QY      257 LeuArgArgLeuAsp 261
Db      844 CTACGTCGTCGAT 858

RESULT 8
AX3464575      1006 bp      DNA      linear      PAT 16-JUL-2002
LOCUS
Sequence 3 from Patent WO0202798.
ACCESSION
AX3464575
VERSION
AX3464575.1 GI:21899370
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
Boyes, D.C., Davis, K.R., Woessner, J.P., Gorchach, J., Hamilton, C.M.,
Hoffman, N.E., Klotz, A.S., Zayed, A., Ascenzi, R.A., Allen, K.,
Mulpuri, R. and Kjemtrup, S.
Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
Patent: WO 0202798-A 3 10-JAN-2002;
Paradigm Genetics Inc. (US)
FEATURES
Location/Qualifiers

```

```
source 1..1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores: 1.73e-70 Length: 1006
Pred. No.: 818.00 Matches: 162
Score: 72.83% Conservative: 31
Percent Similarity: 61.13% Mismatches: 68
Best Local Similarity: 61.09% Indels: 4
Query Match: 6 Gaps: 2
DB:

US-10-624-061-16 (1-261) x AX464575 (1-1006)
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db 64 ATGGCAAGAGTCTTCGAATCGGATTCGGGTCTGTTGTTCCAAATGACTAGTCTTGAC 123
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuLeuGluArg 38
Db 124 TTAATCAGAGAATCGTGTGATAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGA 183
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58
Db 184 GCTAAGTTTCCACTCAATCTCTCGGATTCGAGGAATCTCGTTGCTAGATTCTCGAAGT 243
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 244 TTCTCTTCTCACTAGTCTTTCGTCAGAGACAGACAAATCATCAAGCTAAGGTAGGA 303
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98
Db 304 AGATATGAATACCCGGAAGAGATCTTCTTCTTGAAGAAATTCCTCACTCGGTTTTT 363
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118
Db 364 CTTACGACACAATATCCATCGGCTTTGACCCCTTAAGGCTCTATCTGTTAACTTAACAAA 423
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138
Db 424 CAAATCTGGGATATTACTTTAAAGAAATTCCTTCTTGTGTCAAACCTGCGCGATGAT 483
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArg 158
Db 484 GGCACATATCCATCACTGCTGCTAGTGATCTCGCCTGTTTACAGCTCTTTCGAGAGG 543
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178
Db 544 ATTCACTACGGTAAATTTGATGCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAG 603
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuValLysLeuLeuThrPheThrSerVal 198
Db 604 CTTGCGGATTCGCGCTCAGGATAGAGAGGCTTTGATGAGCTGTGACGTTTGAGAAAGTA 663
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218
Db 664 GAAGAAATGTTAAGAAGAGAGTGCAGAGAAAGAGAAACGTTTGGACAAGAGTAAAG 723
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db 724 TTCAACTCTCGGTATGCGGATGAGAGTAAAGAAAGTATTAAGTGGATCAATGCTTGGC 783
QY 237 SerSerLeuTyrLysAsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeu 256
Db 784 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCT 843
QY 257 LeuArgArgLeuAsp 261
Db 844 CTACGCTGCTCGAT 858

RESULT 9
```

```
AX464581/c
LOCUS AX464581 1006 bp DNA linear PAT 16-JUL-2002
DEFINITION Sequence 9 from Patent WO0202798.
ACCESSION AX464581
VERSION AX464581.1 GI:21899376
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1
AUTHORS Boyes, D.C., Davis, K.R., Woessner, J.P., Gorlach, J., Hamilton, C.M.,
Hoffman, N.E., Kloti, A.S., Zayed, A., Ascenzi, R.A., Allen, K.,
Mulpuri, R. and Kjemtrup, S.
TITLE Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: WO 0202798-A 9 10-JAN-2002;
Paradigm Genetics Inc. (US)
FEATURES
source
1..1006
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Alignment Scores: 1.73e-70 Length: 1006
Pred. No.: 818.00 Matches: 162
Score: 72.83% Conservative: 31
Percent Similarity: 61.13% Mismatches: 68
Best Local Similarity: 61.09% Indels: 4
Query Match: 6 Gaps: 2
DB:

US-10-624-061-16 (1-261) x AX464581 (1-1006)
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db 943 ATGGCAAGAGTCTTCGAATCGGATTCGGGTCTGTTGTTCCAAATGACTAGTCTTGAC 884
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuLeuGluArg 38
Db 883 TTAATCAGAGAATCGTGTGATAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGA 824
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58
Db 823 GCTAAGTTTCCACTCAATCTCTCGGATTCGAGGAATCTCGTTGCTAGATTCTGGAAGT 764
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 763 TTCTCTTCTCACTGAGTTTTTCTGTCAGAGACAGAGAAATCATCCAAGCTTAAAGTAGGA 704
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98
Db 703 AGATATGAATACCCGGAAGAGATCTTCTTCTTGAAGAAATCTCTCACTCGGTTTTT 644
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118
Db 643 CTTACGCAAAATATCCATCGGCTTTCACCCCTTATCTCTGTTAACTTAACATAACAAA 584
QY 119 SerIleTyrLysMetTyrPheLysLeuLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138
Db 583 CAAATCTGGGATATTACTTTAAAGAAATTCCTTCTTGTGTTGTCACACCTCGGCGATGAT 524
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158
Db 523 GGCACATATCCATCACTGCTGCTAGTATCTCGCCTGTTTACAAGCTCTTTCGAGAGG 464
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178
Db 463 ATTCACTACGGTAAATTTGATGCTGAGGTCAAATTCAGAGATGCTCCACAAGATTACGAG 404
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuValLysLeuLeuThrPheThrSerVal 198
```

```

Db 403 CCTGGATTCGGCGCTCAGGATAGAGAGGCTTGATGAAGCTGTGTGACGTTTGAGAAAGTA 344
QY 199 GluGluThrValArgLysArgValGluLysLysLysAlaValPheGlyGlnGluValAsn 218
Db 343 GAAGAAATGTTAAGAGAGAGTGCAGAGAAAGCAGAAAGCTTTGGACAGAGATAAA 284
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db 283 TTCAACTCTGGCTATGGCGATCAGAGAGTAAAGAGAGATATAAAGTGCATCCATTCGTTGCC 224
QY 237 SerSerLeuTyLysAsnTrpValIleProLeuThrLysGluValGlnValGluTrpLeu 256
Db 223 TCTCCGATCTCGGGGAATGGCTTATCCCTCTCACTAAGTCGCTTGAGGTTGAGTATCTT 164
QY 257 LeuArgArgLeuAsp 261
Db 163 CTACGTCGTCGAT 149

```

## RESULT 10

L47355 1006 bp mRNA linear PLN 13-AUG-1999  
 Arabidopsis thaliana chorismate mutase mRNA, complete cds.

DEFINITION L47355  
 ACCESSION L47355  
 VERSION L47355.1 GI:5732015  
 KEYWORDS chorismate mutase; mutase.  
 SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1006)  
 Eberhard J., Ehrler T.T., Epple, P., Felix, G., Raesecke, H.R.,  
 Amrhein, N. and Schmid, J.

Cytosolic and plastidic chorismate mutase isozymes from Arabidopsis  
 thaliana: molecular characterization and enzymatic properties  
 Plant J. 10 (5), 815-821 (1996)

97111372  
 MEDLINE  
 PUBMED 8953244

COMMENT GSDB:S:46595

FEATURES

source

1..1006 location/Qualifiers

/organism="Arabidopsis thaliana"

/mol\_type="mRNA"

/db\_xref="taxon:3702"

/clone="catCM2"

64..861

/EC\_number="5.4.99.5"

/note="putative"

/codon\_start=1

/product="chorismate mutase"

/protein\_id="AAD48922.1"

/db\_xref="GI:5732016"

/translation="MARVFESDSGSGCNVLSLDLIRESLIROEDTIVFSLIERAKFP  
 LNSAPESRCLDSGSFSSLTFFVRETEIIQAKVGRYEVPEPFPFLENIPHSVPT  
 HKYFSALHPKALSNINIKWDIYFKELLPLFVRPGDGNYPSTAASDLACLQALSR  
 IHYKGFVAEVRPAQDYEPAIRAQDEALMKLLTFEKEVMVKRQVKQKAEFTQGE  
 VKFNSGYDESKKRYKVDPLLASRIYGEWLIPLTKUVEVYLLRLD"

1006 /note="putative"

ORIGIN

Alignment Scores:

Pred. No.: 1.73e-70 Length: 1006

Score: 818.00 Matches: 162

Percent Similarity: 72.83% Conservative: 31

Best Local Similarity: 61.13% Mismatches: 68

Query Match: 61.09% Indels: 4

DB: 8 Gaps: 2

US-10-624-061-16 (1-261) x L47355 (1-1006)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyThrLeuAla 18

```

Db 64 ATGCGAAGAGTCTTCCAATCGGATTCGGGTTCGTTGTTCCAATGTACTAGTCTTGAC 123
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleTyLeuLeuLeuGluArg 38
Db 124 TTAATCAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGATCGAGAGA 183
QY 39 AlaLysPheProSerAsnSerHisThrTyAspGluLysTyAlaGlnIleGlnGlyPhe 58
Db 184 GCTAAGTTCCTCAATCTCTCTGCTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGT 243
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 244 TTCTCTCTCTCTACTGAGTTTTCGTCAGAGACAGAAATCATCCAAGCTAAGTAGAGA 303
QY 79 ArgTyLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db 304 AGATATGAATACCCGGAAGAGAAATCTTTCTTCTTCTTGAGAACATTCCTCACCTGGTTT 363
QY 99 ProSerTySerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118
Db 364 CCTACGCCAACAATATCCATCGGCTTTGCACCCCTAAGGCTCTATCTGTTAAACATTACCAA 423
QY 119 SerIleTrpLysMetTyTrpPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138
Db 424 CAAATCTGGATATTTACTTTTAAAGAAATTCCTCTCTTGTGTTGTTCAACCTGGCGATGAT 483
QY 139 GlyAsnTyAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158
Db 484 GGCAACTATCCATCAACTCTCTCTAGTATCTCGGCTGTTTACAAGCTCTTTCGAGAGG 543
QY 159 IleHisTyGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyGlu 178
Db 544 ATTCATAGCTTAATTTGTAGCTGAGTCAAAATTCAGAGATGCTCCACAGATTAACGAG 603
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db 604 CTGCGATTCGGGCTCAGGATAGAGAGCTTTGATGAAGCTGTTGACGCTTGAGAAAGTA 663
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218
Db 664 GAAGAAATGTTAAGAGAGAGTGCAGAGAAAGCAGAAACGTTTGGCAGAGAGTAAAG 723
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236
Db 724 TTCAACTCTCGCTATGGCGATGAGAGTAAAGAAAGTATAAAGTGCATCCATTCGTTGCC 783
QY 237 SerSerLeuTyLysAsnTrpValIleProLeuThrLysGluValGlnValGluTrpLeu 256
Db 784 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843
QY 257 LeuArgArgLeuAsp 261
Db 844 CTACGTCGTCGAT 858

```

## RESULT 11

AY065238

LOCUS

DEFINITION

mRNA, complete cds.

ACCESSION AY065238

VERSION AY065238.1

KEYWORDS FLI CDNA.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 1055)

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,

Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,

Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,

Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,

PLN 18-SEP-2002

Arabidopsis thaliana putative chorismate mutase CM2 (At5g10870)

Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Arabidopsis Full Length cDNA Clones

# REFERENCE

## AUTHORS

2 (bases 1 to 1055)

Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kaniya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.  
Direct Submission  
Submitted (03-DEC-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA  
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kaniya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

# TITLE

## JOURNAL

### COMMENT

The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Banh, J., Chang, C.H., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

## FEATURES

### source

1..1055  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="RAFL09-41-E09 (R18739)"  
/ecotype="Columbia"  
/note="This clone is in a modified pBluescript vector (FIC-1) as a BamHI/XhoI insert."

### gene

1..1055  
/gene="At5g10870"

### 5'UTR

1..98  
/gene="At5g10870"

### misc\_difference

2  
/gene="At5g10870"  
/note="compared to genomic sequence"  
/replace="t"

### CDS

99..896  
/gene="At5g10870"  
/codon\_start=1  
/evidence=experimental  
/product="putative chorismate mutase CM2"  
/protein\_id="P138714.1"  
/db\_xref="GI:17529008"  
/translation="MARVPSDSGSGSNVSLDLRLSRLROEDTIVFSLIERAKFP  
LNSPAPFSCLDGSGSSTFEVRETEIIQAKVGRVYEPENPFLENIPHSVPPT  
HKYPSALHPKALSYNIQWIDYFKELLPFKVPGDGNYPSTAASDLACLQALSR  
IHYGKFAVEKFDAPDYEPIRAODREALMLLTFFKEVEMVKRVQKKAETTFQE  
VKNSGVDSESKKVKYVDPLASRIYGEWLIPTKLVEVEYLLRLD"  
897..1055  
/gene="At5g10870"

### 3'UTR

## ORIGIN

### Alignment Scores:

Pred. No.:	1..83e-70	Length:	1055
Score:	818.00	Matches:	162
Percent Similarity:	72.83%	Conservative:	31
Best Local Similarity:	61.13%	Mismatches:	68
Query Match:	61.09%	Indels:	4
DB:	8	Gaps:	2

US-10-624-061-16 (1-261) x AY065238 (1-1055)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
Db 99 ATGGCAAGAGTCTTCGAATCGGATTCGGTTCCTGTTGTTCCAAATGACTAGTCTTGAC 158  
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
Db 159 TTAATAGAGATTCGTTGATTAGGCAAGACACCATCGTCTTCAGTGTTCAGGATTCAGAGA 218  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
Db 219 GCTAAGTTTCCACTCAATCTCCTGCTTCGAGGATTCCTGTTGCTAGATTCGGAAGT 278  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
Db 279 TTCTCTCTCTCCTCAGTGGTTCCTGCTTCGAGGATTCCTGTTGCTAGATTCGGAAGT 338  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98  
Db 339 AGATATGATATCCCGAGAGAGATTCCTTCTCTTGTGAGACATTCCTCCTCCTCGTGT 398  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
Db 399 CCTAGCGCAATATCCATCGGCTTTGCACCTAAGGCTCTATCTGTTAACAATTAACAAA 458  
QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
Db 459 CAAATCTGGATATTTACTTTAAAGAAATGCTTCTCTTTGTTGTTGTTGTTGTTGTTGTT 518  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
Db 519 GCAACTATCCATCACTGCTGCTAGTATCGCTGTTTACAGCTCTTTTCGAGAAGG 578  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
Db 579 ATTCACTACGGTAAATTTGCTAGCTGAGGTCAAAATTCAGAGATGCTCCCAAGATTACGAG 638  
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
Db 639 CTGCGATTCGCGCTCAGGATAGAGAGGCTTTGATGAAGCTGTTGACGTTTGAGNAAGTA 698  
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValAlaValPheGlyGlnGluValAsn 218  
Db 699 GAAGAAATGTTTACAGAGAGAGTGCAGAAAGACAGAAACGTTTGGACAAAGATAAAA 758  
QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
Db 759 TTCAACTCTGGCTATGGCGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCT 818  
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
Db 819 TCTGCGATCTACGGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCT 878  
QY 257 LeuArgArgLeuAsp 261  
Db 879 CTACGCTGCTCGAT 893

## RESULT 12

### AR404683

### LOCUS

### DEFINITION

### ACCESSION

### VERSION

### KEYWORDS

### SOURCE

AR404683 Sequence 5 from patent US 6627798.  
AR404683  
AR404683.1 GI:40153351  
Unknown.  
525 bp DNA linear PAT 18-DEC-2003

ORGANISM Unknown.  
Unclassified.  
REFERENCE 1 (bases 1 to 525)  
AUTHORS Falco, S.C., Farnoud, O.O. and Lee, J.-M.  
TITLE Aromatic amino acid biosynthetic enzymes  
JOURNAL Patent: US 6627799-A 5 30-SEP-2003;  
FEATURES Location/Qualifiers  
source 1..525  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 6,41e-66 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.29% Conservative: 0  
Best Local Similarity: 95.29% Mismatches: 7  
Query Match: 57.32% Indels: 4  
DB: 6 Gaps: 0

US-10-624-061-16 (1-261) x AR404683 (1-525)  
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
DB 16 ATGGCCAAAGCAGCAGACAAAGTCTTGATTCCTGGGATGTGTACACCGTAGCTTCGTG 75  
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
DB 76 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGGTCTCATTTGAGAGAGCCAA 135  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
DB 136 TTCCTTAGCAATCTCACACCTATGATGAAAGATGATCTCAATCCAGGGTTTTGTGGC 195  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
DB 196 TCATTGGTGGATTTGTTGTTAAGAATACAGAGGCCATTCAGCTTAAGGCTGAAGATAC 255  
QY 81 LysAsnProGluGluAsnAlaPheProGluAsnLeuProSerIleValProSer 100  
DB 256 AAAACCCCTGAAGAAACCCCTCTTCCAGAAATTTTACACCATCAATTTGTGCCATCT 315  
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
DB 316 TACTCTCTCAACAGCTTTTGTGATCTCTGCTGCTTCAATTAACATAACAAGTCATCT 375  
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
DB 376 GGGAAATGATTTTCAAGAGGTACTTCCATTGCTTGTCTTCCGGGTGATGAGNAAC 435  
QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHi 160  
DB 436 TATGGCAAACTGACGCTTAAGACCT-TCATTTATGACGTC-ATCTCTAGAGGAT-CA 492  
QY 160 sTyrGlyLysPheValAlaGluValLys 169  
DB 493 CTATGAAAGTTGTGACTGANGGAA 520

RESULT 13  
LOCUS AX506042 1005 bp DNA linear PAT 27-SEP-2002  
DEFINITION Sequence 737 from Patent WO0216655.  
ACCESSION AX506042  
VERSION AX506042.1 GI:23387279  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1  
REFERENCE Harper, J.F., Kreps, J., Wang, X. and Zhu, T.  
AUTHORS Stress-regulated genes of plants, transgenic plants containing

same, and methods of use  
Patent: WO 0216655-A 737 28-FEB-2002;  
The Scripps Research Institute (US) ; Syngenta Participations AG  
(CH)  
FEATURES Location/Qualifiers  
source 1..1005  
/organism="Arabidopsis thaliana"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:3702"

ORIGIN  
Alignment Scores:  
Pred. No.: 8,81e-59 Length: 1005  
Score: 898.00 Matches: 131  
Percent Similarity: 68.08% Conservative: 46  
Best Local Similarity: 50.38% Mismatches: 75  
Query Match: 52.13% Indels: 8  
DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x AX506042 (1-1005)  
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnLys 29  
DB 223 GATGAGAGTTCAGAGTTGACTCTTGAAGGTATTAGAACTCTTTGATCGGTCAAGAGGAC 282  
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
DB 283 AGCATATATTTGGGCTATTGGAGAGCCAGTACTGTACATGCTCATACTTATGAT 342  
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
DB 343 CCTACTGCTTTTGACATGATGTTTCAATGGTCTCTTGGTTGAGTACATGTTAAAGGC 402  
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
DB 403 ACTGGAAGCTTCACCTAAGTGTGGTAGTTTAAAGTCTTGAAGTCTGATGACATCTTCTT 462  
QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
DB 463 CCTGATGATCTACAGAGCTATGTTGCTCTCTTCAGTACCCAAAGGTGTGCATTT 522  
QY 110 GlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
DB 523 GCTGCTGATTCGATAAACAATAACAAGAGATATGGAATGATCTTACAGACCTTTGT 582  
QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
DB 583 CCAGACTTGTGAGAAAGCGGATGATGTTAACTACGGCTCAACAGCTCTCTGTGAGCT 642  
QY 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
DB 643 ATCTGCTTTCAGTGTCTCTCAAGAGAGATCCATTAACGGTAAATTTGTTCCAGAGCTAAA 702  
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
DB 703 TTTCAAGCCCTCACCCGAGCATACGATCGGCTCAAGCAAGATGAAGATGAGTGCATG 762  
QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
DB 763 ATGGATATGTCGACATTCGCCGACTGTGGAAGATCGCATTAAGAAAGAGAGTTCAGATGAAA 822  
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAspAsnGluAsnArg 229  
DB 823 ACCCGAATACACGGGCAAGAGAGTGGATGGAGGAGAAAGAGAGAGAGAGTTCAGATGAAA 882  
QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241  
DB 883 GAAGGGAATGATCTCATGTTTACAAAATCAGTCCGATCTTAGTTGGTGACATATATGGA 942  
QY 242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261  
DB 943 GATTGGATCATGCTTTTAAACAAAAGAGGTTCAGTGGAGTACTTGTCTCAGAGAGCTGGAC 1002

```
RESULT 14
AR236633
LOCUS AR236633 1207 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 2 from patent US 6465217.
ACCESSION AR236633
VERSION AR236633.1 GI:27280719
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1207)
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.
TITLE Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: US 6465217-A 2 15-OCT-2002;
FEATURES
LOCATION/Qualifiers
SOURCE 1..1207
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,74e-58 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 88.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.98% Indels: 8
Gaps: 6
DB: 1
US-10-624-061-16 (1-261) x AR236633 (1-1207)
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
DB 235 GATCAGAGTGGAGTTTGACTCTTGAAGGTATTAGAACTCTTTGATCCGTCAGAGGAC 294
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
DB 295 AGCATTATATTGGGCTATTGGAGAGCCCAAGTACTGTACAACTGCTGATGATGAT 354
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
DB 355 CCTACTGCTTTTCAGATGGTGGTTTCAATGCTTCTTGGTTGAGTACATGTTAAAGGC 414
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
DB 415 ACTGAGAAGCTTCACGCTAAGTTGGTAGTTTAAGAGTCCCTGATGAACATCCTTTCTTC 474
QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
DB 475 CCTGATGATCTACAGAGCCTATGTTGCTCTCTTCAGTACCCAAAGGTGTTGCAATTT 534
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
DB 535 GCTGCTGATTCGATAACATAACAAGAAGATATGAACATGATGATTCAGAGACCTTGT 594
QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
DB 595 CCAAGACTTGTGAAGAAGCCGATGATGTAACATACGGCTCAACAGCTGTCTGTGACGT 654
QY 150 SerLeuLeuGlnSerIleSerArgGileHisTyrGlyLysPheValAlaGluValLys 169
DB 655 ATCTGCCCTTCAGTGTCTCAAGAGAGATCCATTACGGTAAATTTGTCAGAGAGCTAA 714
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
DB 715 TTTCAAGCCTCACCCGAAGCATACGAGTCGCCCATCAAGACCAAGATAAGGATCGACTG 774
QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
DB 775 ATGATATGCTGACATTCCTCCGACTGTGGAAGATGCCGATAAAGAAGAGAGTTGAGATGAA 834
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAspAsnGluAsnArg 229
DB 835 ACCCGACATACGGCAAGAAGTCAAGTTGGGATGGAGCAAGAAGAAGAAGAA 894
QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
DB 895 GAAGGGAATGAATCTCATGTTTACAAAATCAGTCCGATCTTAGTTGGTGAATATGGA 954
QY 242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuValArgLeuAsp 261
DB 955 GATTGGATCATGCTTTTACAAAAGAGGTTCAAGTGGAGTACTTCTCAGAAGACTGAC 1014
RESULT 15
AR236639/c
LOCUS AR236639 1207 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 8 from patent US 6465217.
ACCESSION AR236639
VERSION AR236639.1 GI:27280725
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1207)
AUTHORS Boyes,D.C., Davis,K.R., Woessner,J.P., Goriach,J., Hamilton,C.M.,
Hoffman,N.E., Kloti,A.S., Zayed,A. and Ascenzi,R.A.
TITLE Methods and compositions for the modulation of chorismate synthase
and chorismate mutase expression or activity in plants
JOURNAL Patent: US 6465217-A 8 15-OCT-2002;
FEATURES
LOCATION/Qualifiers
SOURCE 1..1207
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1,74e-58 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 88.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.98% Indels: 8
Gaps: 6
DB: 1
US-10-624-061-16 (1-261) x AR236639 (1-1207)
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
DB 973 GATCAGAGTGGAGTTTGACTCTTGAAGGTATTAGAACTCTTTGATCCGTCAGAGGAC 914
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
DB 913 AGCATTATATTGGGCTATTGGAGAGCCCAAGTACTGTACATGCTGATGATGAT 854
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
DB 853 CCTACTGCTTTTGACATGATGCTGTTTCAATGCTTCTTGGTTGAGTACATGTTAAAGGC 794
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
DB 793 ACTGAGAAGCTTCACGCTAAGTTGGTAGTTTAAGAGTCCCTGATGAACATCCTTTCTTC 734
QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
DB 733 CCTGATGATCTACAGAGCCTATGTTGCTCTCTTCAGTACCCAAAGGTGTTGCAATTT 674
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
DB 673 GCTGCTGATTCGATAACATAACAAGAAGATATGGAACATGATGATCTTCAAGAGACTGTT 614
QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
DB 613 CCAAGACTTGTGAAGAAGCCGATGATGTAACATCGCTCAACAGCTGTCTGTGACGCT 554
QY 150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
DB 150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
```



```
Db      553 ATCTGCCTTCACTGCTCTCTCAAAGAGAAATCCATTACGGTAAATTTGTTGCAGAAAGCTAAA 494
Qy      170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      493 TTTCAAGCCTCACCCGAGCATACGAGTCGGCCATCAAAGCACAGATAGATCGACTG 434
Qy      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      433 ATGGATATGCTCACATTCCTCGACTCTGGAAGATGGGATAAAGAGAGAGATTGAGATGAAA 374
Qy      210 AlaValValPheGlyGluGluValAsnLeuAsnSerAspAspAsnAspAsnGluAsnArg 229
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      373 ACCCGAACATACGGGCAAGAGTGAAGTTGGGATGGAGGAGAAAGAAAGAAAGAGAGAA 314
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      313 GAAGGGAAATGAATCTCATGTTTACAAATCAGTCCGATCTTAGTTGGTGACTTATATGGA 254
Qy      242 AsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
      |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      253 GATTGGATCATGCCTTTAAACAAAGAGGTTCAAGTGGAGTACTTCTCAGAAAGACTGGAC 194
```

Search completed: September 25, 2004, 04:17:47  
Job time : 4199 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:54:52 ; Search time 433 Seconds  
(without alignments)

2560.692 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQSPDGNVYTLASV.....NWVPLTKEVQVEYLLRLRD 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgn2\_1/USPTO.spool/US10624061/runat\_24092004\_092406\_16902/app\_query.fasta\_1.455

-DB=N Geneseq\_29Jan04 -OFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blcsum62 -TRANS=human40.cdi

-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US10624061@cgn\_1\_1\_470 @runat\_24092004\_092406\_16902 -NCPU=6 -ICPU=3

-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Geneseq\_29Jan04.\*

2: Geneseq\_1980s.\*

3: Geneseq\_1990s.\*

4: Geneseq\_2000s.\*

5: Geneseq\_2001as.\*

6: Geneseq\_2001bs.\*

7: Geneseq\_2002s.\*

8: Geneseq\_2003as.\*

9: Geneseq\_2003bs.\*

10: Geneseq\_2003cs.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1339	100.0	1020	7 ABX93058	Abx93058 Soybean c
2	818	61.1	798	6 ABZ13167	Abz13167 Arabidops
3	818	61.1	993	6 ABA99614	Aba99614 Construct
4	818	61.1	1006	6 ABA99613	Aba99613 A. thalia
5	818	61.1	1006	6 ABA91400	Abag1400 Arabidops
6	818	61.1	1006	6 ABA91394	Abag1394 Arabidops
7	767.5	57.3	525	7 ABX93053	Abx93053 Soybean c
8	698	52.1	1005	6 ABZ12932	Abz12932 Arabidops

C	9	696	52.0	1207	6	ABA91393	ABA91393 Arabidops
	10	696	52.0	1207	6	ABA91399	ABA91399 Arabidops
	11	685.5	51.2	1231	7	ABX93057	Abx93057 Rice chor
	12	677.5	50.6	1274	3	AAC33324	Aac33324 Arabidops
C	13	670.5	50.1	1223	7	ABX93056	Abx93056 Corn chor
	14	654.5	48.9	1217	6	ABA91395	Abag1395 Arabidops
	15	654.5	48.9	1217	6	ABA91401	Abag1401 Arabidops
	16	602	45.0	780	7	ABX93059	Abx93059 Wheat cho
C	17	587	43.8	5176	6	ABA91397	Abag1397 Arabidops
	18	549.5	41.0	5099	6	AAD23643	Aad23643 Arabidops
	19	549.5	41.0	5099	6	ABA91396	Abag1396 Arabidops
	20	522	39.0	579	7	ABX93054	Abx93054 Wheat cho
C	21	423	31.6	843	3	AAC81949	Aac81949 H. polyom
	22	423	31.6	1655	3	AAC81950	Aac81950 H. polyom
	23	402.5	30.1	869	7	AAL50202	Aal50202 M sterilli
	24	371.5	27.7	2920	7	ABT17891	Abt17891 Aspergilli
C	25	371.5	27.7	2987	7	ABT19705	Abt19705 Aspergilli
	26	312.5	23.3	658	3	AAF14469	Aaf14469 Aspergilli
	27	267	19.9	287	6	ABL72637	AbL72637 Corn tass
	28	248	18.5	258	6	ABF71526	Abf71526 Corn tass
C	29	144.5	10.8	560	4	AAH44222	Aah44222 Physcomit
	30	144.5	10.8	601	7	ABX93052	Abx93052 Rice chor
	31	130	9.7	552	7	ABX93051	Abx93051 Corn chor
	32	121.5	9.1	1379	6	ABA91794	Abag1794 Yeast mit
C	33	113	8.4	507	3	AAF07985	Aaf07985 Fusarium
	34	105	7.8	1080	7	ABX05983	Abx05983 S. pneumo
	35	105	7.8	19250	2	AAV52168	Aav52168 Streptoco
	36	105	7.8	110000	7	ASB56454	ASB56454 <sup>102</sup> Streptoc
C	37	103	7.7	2346	4	AAS53535	Aas53535 Haemophil
	38	103	7.7	2346	7	ACA34520	Aca34520 Prokaryot
	39	103	7.7	110000	2	AAT42063	Continuation (18 o
	40	100.5	7.5	2385	9	ADD30303	Add30303 Plant yie
C	41	96.5	7.2	2043	6	ABZ13532	Abz13532 Arabidops
	42	96.5	7.2	2043	8	ADB23149	Adb23149 Enviroame
	43	96	7.2	1380	4	AAH53580	Aah53580 S. epider
	44	96	7.2	1380	4	AAH52926	Aah52926 S. epider
45	96	7.2	1395	6	ABN92401	Abn92401 Staphyloc	

#### ALIGNMENTS

##### RESULT 1

ABX93058

ID ABX93058 standard; cDNA; 1020 BP.

XX AC ABX93058;

XX DT 22-MAY-2003 (first entry)

XX DE Soybean chorismate mutase cDNA.

XX KW Soybean; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;  
XX KW biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;  
XX KW phenylalanine; prephenate dehydrogenase; herbicide; food crop.

XX OS Glycine max.

XX FH Key Location/Qualifiers

XX FT CDS 23..808

XX FT /\*tag= a

XX FT /product= "Chorismate mutase"

XX PN US2002184659-A1.

XX PD 05-DEC-2002.

XX PF 03-DEC-1999; 99US-00454279.

XX PR 04-DEC-1998; 98US-0110845P.

XX PA (FALC/) FALCO S C.

XX PA (FAMO/) FAMODU O O.

XX PA (LEEJ/) LEE J.

XX Falco SC, Famodu OO, Lee J;  
 XX WPI; 2003-328651/31.  
 DR P-PSDB; ABU08094.  
 XX  
 PT New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
 PT nucleic acids, useful for facilitating design and/or identifying  
 PT inhibitors of those enzymes that may be used as herbicides and for  
 PT producing antibodies.  
 XX  
 PS Claim 3; Page 22; 32pp; English.  
 XX  
 CC The invention discloses isolated polynucleotides encoding chorismate  
 CC mutase polypeptides. Aromatic amino acids must be included in the diets  
 CC of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
 CC converted to anthranilate during tryptophan biosynthesis and is converted  
 CC to prephenate, the branch point for tyrosine and phenylalanine  
 CC biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
 CC prephenate. Also disclosed are methods for selecting an isolated  
 CC polynucleotide that affects the level of expression of a tyrosine  
 CC biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
 CC fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
 CC evaluating compounds for their ability to inhibit the activity of a  
 CC tyrosine biosynthetic enzyme. The polypeptides can be used produce  
 CC antibodies. Chorismate mutase and prephenate dehydrogenase are good  
 CC targets for herbicides that will not affect animals, and overexpression  
 CC of these enzymes may be used to increase the content of aromatic amino  
 CC acid in food crops. The polypeptides may also be used to design and/or  
 CC identify inhibitors of those enzymes that may be used as herbicides. The  
 CC nucleic acids may be used to create transgenic plants, as probes for the  
 CC genetic and physical mapping of the genes and as markers for traits  
 CC linked to those genes. The sequence presented is a soybean chorismate  
 CC mutase cDNA  
 XX  
 SQ Sequence 1020 BP; 322 A; 176 C; 223 G; 299 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,32e-148 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-624-061-16 (1-261) x ABX93058 (1-1020)

QY 1 MetalAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValThrThrLeuAlaSerVal 20  
 DB 23 ATGGCCAAAGCAGCAGACAAAGTCTGATTCGGGAATGTTACAGCTAGCTTCGTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleLeuGlyLeuIleGluArgAlaLys 40  
 DB 83 AGAGAGGATTGGTTAGCAAGAGGATACCATCTATTATGCTCTCATTTAGAGAGCCAAAG 142  
 QY 41 PheProSerAsnSerHisThrTyraSpGluLysTyraAlaGlnIleGlnGlyPheCysGly 60  
 DB 143 TTCCTAGCAATTCACACCATATGATGANAAGATGCTCAATCCAGGGTTTGTGGC 202  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 DB 203 TCATTGGTGGATTTGTTTAAAGATACAGAGCCATTCAGCTAGGCTGGAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 263 AAAAACCCCTGAAGAAACGCCCTCTCTCCAGAAAATTTTACCACCATCAATTGTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyValAlaSerIleAsnLysSerIle 120  
 DB 323 TACTCTTCAAAACAGTTTTTCATCCCTGGTCTGCTCAATTACATTAACACAGTCCATC 382  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140

DB 383 TGGAAATGTATTTCACAAAGAGTTACTTCATTGCTGCTTACCTTCGGGTGATGATGCAAC 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160  
 DB 443 TATGCGCAACTGCAGCTAATGACCTTTTCATTATTGCACTCCATCTCTAGAAAGGATTAC 502  
 QY 161 TyrGlyPhePheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 503 TATGGAAGTTTGTAGCTAGGTTGAAATTCAGGGATGCTCTCAAGACTACGAGCTTTA 562  
 QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
 DB 563 ATTCGAGCTAAGATTAAGAGGATTGATGAATTTTGACATTTACAGCGTTGAGAG 622  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 DB 623 ACGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTGTTTGGGCAGGAGTGAATCTTAAC 682  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
 DB 683 AGTCATGACATGACATTAAGAAACCGTAATTTGATCCATCATGCTTCTAGCTTGTAC 742  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeu 260  
 DB 743 AAAAATGGGTGATACCTCTCACCAGGAGGTTTCAAGTTGAGTACCTCTTGGCGGTCTA 802  
 QY 261 Asp 261  
 DB 803 GAC 805  
 RESULT 2  
 ABZ13167  
 ID ABZ13167 standard; DNA; 798 BP.  
 XX  
 AC ABZ13167;  
 XX  
 DT 21-JAN-2003 (first entry)  
 XX  
 DE Arabidopsis thaliana stress regulated gene SEQ ID NO 972.  
 XX  
 KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN WO200216555-A2.  
 XX  
 PD 28-FEB-2002.  
 XX  
 PF 24-AUG-2001; 2001WO-US026685.  
 XX  
 PR 24-AUG-2000; 2000US-0227866P.  
 PR 25-JAN-2001; 2001US-0264647P.  
 PR 22-JUN-2001; 2001US-0300111P.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
 XX  
 PI Harper JF, Krepis J, Wang X, Zhu T;  
 XX  
 WPI; 2002-304127/34.  
 XX  
 PT Identifying a stress condition to which a plant cell has been exposed and  
 PT producing plants with increased tolerance to these abiotic stresses.  
 XX  
 PS Claim 144; SEQ ID NO 972; 577pp + Sequence Listing; English.  
 CC The invention relates to identifying a stress condition to which a plant  
 CC cell has been exposed, comprising: (a) contacting nucleic acid  
 CC representative of expressed polynucleotides in the plant cell with an  
 CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants

CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 XX Sequence 798 BP; 223 A; 162 C; 180 G; 233 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 5.09e-87 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservativity: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x ABZ13167 (1-798)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 Db 1 ATGGCAAGAGTCTCGAATCGGATTCGGGTTCGGTTCGATGTTCCAAATGACTAGTCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleIleIleIleIleIleIle 38  
 Db 61 TTAATCAGAGAATCGTGTGATTAGGCAAGAAGACACCATCGTCTTCAGCTTGATCGAGAGA 120  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 Db 121 GCTAAGTTTCCACTCAATCTCTGTTTCGAGGAATCTGTTGTTCTAGATTCTGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 Db 181 TTCCTCTCTCACTAGTCTTTCGTCAGAGACAGAAATCATCAAGCTAAGGTAGA 240  
 QY 79 ArgTyrLysAsnProGluGlnAlaPhePheProGluAsnLeuProSerIleVal 98  
 Db 241 AGATATGAATACCCGGAAGAGATCTCTTCTTCGAGAAATCTCTCCTCAGTCTGTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 Db 301 CTTACGCACAAATATCCATCGGCTTTGACCTTACGCTCTATCTGTTTAACTTAACAA 360  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
 Db 361 CAAATCTGGGATATTTACTTAAAGAAATGCTTCTCTTGTGTCAAACCTGCGCATGAT 420  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
 Db 421 GGCACATATCCATCACTGCTGATGATCTCGCTGTTTACAGCTCTTTTCGAGAAGG 480  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 Db 481 ATTCACATCGGTAATTTGTAGCTGAGGTCAAATTCAGAGATGCTCCACAGATTACGAG 540  
 QY 179 ProIleLeuArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 Db 541 CCTCGGATTCGCGCTCAGGATAGAGAGCTTTGATGAGCTGTGACGTTTGAGAAAGTA 600  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218  
 Db 601 GAAGAAATGTTAAGAGAGAGTGCAGAGAAGACAGAAACGTTTCGACAGAGAGTAAAA 660  
 QY 219 LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 Db 661 TTCAACTCTGGCTATGCGGATGAGAGTAAAGAGAGTAAAGTGGATCAATGCTTGGCT 720  
 QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 Db 721 TCTCGGATCTACCGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCT 780  
 QY 257 LeuArgArgLeuAsp 261  
 Db 781 CTACGTCGTCCTCGAT 795

## RESULT 3

ABA99614  
 ID ABA99614 standard; DNA; 993 BP.  
 AC ABA99614;  
 XX 17-MAY-2002 (first entry)  
 DT Construct PCR-script/AtCM-LTP-AtCM2 DNA.

DE Shikimate pathway; chorismate mutase; gene; prephenate dehydrogenase;  
 KW vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;  
 KW feed; pharmaceutical; cosmetic; aromatic amino acid; salicylic acid;  
 KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;  
 KW tocopherol; tocotrienol; ds.

OS Arabidopsis thaliana.

OS Synthetic.

PN WO200200901-A1.

PD 03-JAN-2002.

PF 28-JUN-2001; 2001WO-EP007391.

PR 29-JUN-2000; 2000DE-01030647.

PR 21-DEC-2000; 2000DE-01064454.

PA (SUNG-) SUNGENE GMBH & CO KGAA.

XX Badur R, Geiger M, Kunze I, Sommer S;

XX WPI; 2002-164442/21.

XX P-PSDB; AAM49655.

XX Preparing fine chemicals, particularly Vitamins E and K, useful as  
 PT antioxidants e.g. in foods or medicine, by growing organisms with altered  
 PT shikimate biosynthesis pathway.

XX Claim 11; Page 62-63; 74pp; German.

XX This invention describes a novel method for the preparation of fine  
 CC chemicals by culturing organisms in which the shikimate pathway has been  
 CC altered relative to the wild type. The method involves a construct  
 CC containing a plastid transit peptide, a chorismate mutase and/or  
 CC prephenate dehydrogenase, linked to at least one regulatory sequence for  
 CC transcription or translation in plants. The method is used to produce  
 CC chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
 CC fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
 CC also aromatic amino acids, salicylic or folic acid derivatives,  
 CC phenylpropanoids, flavanoids etc., especially in transgenic plants. Also  
 CC plants with increased Vitamin E contents have improved resistance to  
 CC abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
 CC shikimate pathway are useful as foods, fodder and in preparation of  
 CC processed foodstuffs. Transgenic plants with modified shikimate pathways  
 CC have increased content of desired chemicals, particularly tocopherols  
 CC and/or tocotrienols. This sequence encodes the construct PCR-Script/AtCM-  
 CC LTP-AtCM-2 which is composed of the Arabidopsis thaliana plastid  
 CC chorismate mutase-1 transit peptide and a Arabidopsis thaliana chorismate  
 CC mutase-2 fragment and is used for plastid localisation

XX Sequence 993 BP; 259 A; 227 C; 210 G; 297 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 6.93e-87 Length: 993  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservativity: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 6 Gaps: 2

US-10-624-061-16 (1-261) x ABA99614 (1-993)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 DB 196 ATGGCAAGAGCTTCGAATCGGATTCGGTTCGGTTCCTCAATGACTAGTCTTGAC 255  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 DB 256 TTAATCAGAGAATCGTTGATTAGGCAAGAACACCATCGTCTTCAGCTTGATCGAGAGA 315  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 DB 316 GCTAAGTTCCACTCAATCTCCTGCTTCAGAGATCTCGTCTAGATTCTGGAGT 375  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 DB 376 TTCCTCTCTCACTGAGTTTTCGTCAGAGACAGAAATCATCAAGCTTAAGTAGGA 435  
 QY 79 ArgTyrLysAsnProGluGlnAsnAlaPhePheProGluLysLeuProSerIleVal 98  
 DB 436 AGATATGAATACCCGAGAGAGATCTTCTCTTGAAGACATCTCTCACTCGGTTTT 495  
 QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
 DB 496 CTTACGCAAAATATCATTCATCGCTTTCACCTTAAGGCTCTATCTGTTAACTAACA 555  
 QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 DB 556 CAAATCTGGATATTACTTTAAAGAAATGCTTCTTGTTCAAACCTGGCGATGAT 615  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArg 158  
 DB 616 GCAACTATCATCATCTGCTGCTAGTCTCTCGCTTTTACAGCTCTTTCGAGAGG 675  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 DB 676 ATTCACTACGGTAAATTTGAGTCTGAGGTCAAAATTCAGAGATGCTCCACAGATTAC 735  
 QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198  
 DB 736 CTGGATTCGGCTCAGCAGATAGAGGCTTTGATGAAGCTGTGAGCTTTGAGAAAGTA 795  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValAlaPheGlyGlnGluValAsn 218  
 DB 796 GAAGAAATGCTTAAAGAGAGAGTGCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 855  
 QY 219 LeuAsnSerAspAspAsnAsnGluAsnArgLysPhe-----AspProSerValAla 236  
 DB 856 TTCAACTCTGGCTATGGCGATGAGAGTAAAGAGAGATATAAGTGGATCTTCTGCC 915  
 QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 DB 916 TCTCGACTCATCGGGAATGGCTTATCTCTCTCACTAGCTCGTGTGAGTTGAGTATCTT 975  
 QY 257 LeuArgArgLeuAsp 261  
 DB 976 CTACGCTGCTCGAT 990

RESULT 4  
 ID ABA99613 standard; DNA; 1006 BP.  
 XX AC ABA99613;  
 XX 17-MAY-2002 (first entry)  
 XX DE A. thaliana cytosolic chorismate mutase-2 DNA.  
 XX Shikimate pathway; chorismate mutase; gene; prephenate dehydrogenase;  
 KW vitamin E; vitamin K; ubiquinone; fat-soluble antioxidant; transgenic;  
 KW feed; pharmaceutical; cosmetic; aromatic amino acid; salicylic acid;  
 KW folic acid; phenylpropanoid; flavanoid; plant; abiotic stress resistance;  
 KW tocopherol; tocotrienol; ds.  
 XX

OS Arabidopsis thaliana.  
 XX Key Location/Qualifiers  
 FT 64..861  
 FT CDS /tag= a  
 FT /product= "chorismate mutase-2"  
 XX  
 XX WO200200901-A1.  
 XX 03-JAN-2002.  
 XX 28-JUN-2001; 2001WO-EP007391.  
 XX 29-JUN-2000; 2000DE-01030647.  
 XX 21-DEC-2000; 2000DE-01064454.  
 XX (SUNG-) SUNGENE GMBH & CO KGAA.  
 XX Badur R, Geiger M, Kunze I, Sommer S;  
 XX WPI; 2002-164442/21.  
 XX P-PSDB; AAM49654.  
 XX  
 XX Preparing fine chemicals, particularly Vitamins E and K, useful as  
 PT antioxidants e.g. in foods or medicine, by growing organisms with altered  
 PT shikimate biosynthesis pathway.  
 XX  
 XX Disclosure; Page 57-60; 74pp; German.  
 XX  
 XX This invention describes a novel method for the preparation of fine  
 CC chemicals by culturing organisms in which the shikimate pathway has been  
 CC altered relative to the wild type. The method involves a construct  
 CC containing a plasmid transit peptide, a chorismate mutase and/or  
 CC prephenate dehydrogenase, linked to at least one regulatory sequence for  
 CC transcription or translation in plants. The method is used to produce  
 CC chemicals, particularly Vitamins E or K and/or ubiquinone, useful e.g. as  
 CC fat-soluble antioxidants (in feeds, pharmaceuticals and cosmetics), but  
 CC also aromatic amino acids, salicylic or folic acid derivatives,  
 CC phenylpropanoids, flavanoids etc., especially in transgenic plants. Also  
 CC plants with increased Vitamin E contents have improved resistance to  
 CC abiotic stress, e.g. frost and drought. Transgenic plants with an altered  
 CC shikimate pathway are useful as foods, fodder and in preparation of  
 CC processed foodstuffs. Transgenic plants with modified shikimate pathways  
 CC have increased content of desired chemicals, particularly tocopherols  
 CC and/or tocotrienols. This sequence encodes the Arabidopsis thaliana  
 CC cytosolic chorismate mutase-2 protein used in the method of the invention  
 XX  
 XX Sequence 1006 BP; 290 A; 188 C; 226 G; 302 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 7,06e-87 Length: 1006  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: Gaps: 2  
 US-10-624-061-16 (1-261) x ABA99613 (1-1006)  
 QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 DB 64 ATGGCAAGAGCTTCGAATCGGATTCGGTTCGGTTCCTCAATGACTAGTCTTGAC 123  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 DB 256 TTAATCAGAGAATCGTTGATTAGGCAAGAACACCATCGTCTTCAGCTTGATCGAGAGA 183  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 DB 316 GCTAAGTTCCACTCAATCTCCTGCTTCAGAGATCTCGTCTAGATTCTGGAGT 243  
 QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78





KW aromatic amino acid; diet; biosynthesis; chorismate; anthranilate;  
KW tryptophan; prephenate; tyrosine; phenylalanine;  
KW prephenate dehydrogenase; herbicide; food crop.  
XX  
OS Glycine max.

Key Location/Qualifiers  
CDS 43..438  
/\*tag= a  
/product= "Chorismate mutase"  
/partial  
/note= "No start or stop codon shown"

US2002184658-A1.  
05-DEC-2002.  
03-DEC-1999; 99US-00454279.  
04-DEC-1998; 98US-0110845P.  
(FALCO) FALCO S C.  
(FALCO) FAMODU O O.  
(LEEJ) LEE J.  
Falco SC, Famodu OO, Lee J;  
WPI; 2003-328651/31.  
P-PSDB; AB008089.  
New tyrosine biosynthetic enzyme, chorismate dismutase, proteins and  
nucleic acids, useful for facilitating design and/or identifying  
inhibitors of those enzymes that may be used as herbicides and for  
producing antibodies.

Claim 3; Page 16; 32pp; English.  
The invention discloses isolated polynucleotides encoding chorismate  
mutase polypeptides. Aromatic amino acids must be included in the diets  
of animals. In the aromatic amino acid biosynthetic pathway chorismate is  
converted to anthranilate during tryptophan biosynthesis and is converted  
to prephenate, the branch point for tyrosine and phenylalanine  
biosynthesis. Chorismate mutase catalyses the conversion of chorismate to  
prephenate. Also disclosed are methods for selecting an isolated  
polynucleotide that affects the level of expression of a tyrosine  
biosynthetic enzyme polypeptide in a host cell, obtaining a nucleic acid  
fragment encoding a tyrosine biosynthetic enzyme polypeptide and  
evaluating compounds for their ability to inhibit the activity of a  
tyrosine biosynthetic enzyme. The polypeptides can be used produce  
antibodies. Chorismate mutase and prephenate dehydrogenase are good  
targets for herbicides that will not affect animals, and overexpression  
of these enzymes may be used to increase the content of aromatic amino  
acid in food crops. The polypeptides may also be used to design and/or  
identify inhibitors of those enzymes that may be used as herbicides. The  
nucleic acids may be used to create transgenic plants, as probes for the  
genetic and physical mapping of the genes and as markers for traits  
linked to those genes. The sequence presented is a soybean chorismate  
mutase expressed sequence tag (EST)

Sequence 525 BP; 160 A; 103 C; 111 G; 148 T; 0 U; 3 Other;  
Alignment Scores:  
Pred. No.: 2,72e-81 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.29% Conservative: 0  
Best Local Similarity: 95.29% Mismatches: 7  
Query Match: 57.32% Indels: 4  
DB: 7 Gaps: 0

US-10-624-061-16 (1-261) x ABX93053 (1-525)  
Qy 1 MetAlaLysAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
|||||

Db 16 ATGCCCAAGCAGCAGACACAAAGTCTGATTCTGGGAATGTGTACACGCTAGCTTTCTGTG 75  
Qy 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
Db 76 AGAGAGGATTGGTTAGGCAAGAGGATACCACTTAITTAIGGTCTCATTTAGAGAGGCAAG 135  
Qy 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
Db 136 TTCCCTAGCAATTTCTCACACCTATGATGAAGAAGTATGCTCAATCCAGGGTTTTTGTGGC 195  
Qy 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
Db 196 TCATTGGTGGAAATTTGTTTGAAGATAGAGAGCCATTCAGCTAAGCTGAGCTGGAAGATAC 255  
Qy 81 LysAsnProGluGluAsnAlaPhePheProGluLysAsnLeuProSerIleValProSer 100  
Db 256 AAAAAACCTGAGAAAAAGCCCTTTCCACAGAAAAATTTACCACTCAATTTGCGCATCT 315  
Qy 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
Db 316 TACTCCTTCAACACAGTTTTTGCATCTCTGGTGGCTTCAATTAACATAACAAGTCACTCT 375  
Qy 121 TrpLysMetTyrPheLysGluLeuLeuProLeuAlaThrSerGlyAspGlyAsn 140  
Db 376 GGGAAATGATTTTCAAGAGATTACTTCCATTGCTTGTCTTCTCGGTGATGATGNAAC 435  
Qy 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHi 160  
Db 436 TATGCGCAAACTGCAGCTAATGACCT-TCAATTTATTCAGTTC-ATCTCTAGAAGGAT-CA 492  
Qy 160 STYRGlyLysPheValAlaGluValLys 169  
Db 493 CTATGGAAGTTTGTAGCTGANGNGAAA 520  
RESULT 8  
ABZ12932  
ID ABZ12932 standard; DNA; 1005 BP.  
XX  
AC ABZ12932;  
XX  
DT 21-JAN-2003 (first entry)  
XX  
DE Arabidopsis thaliana stress regulated gene SEQ ID NO 737.  
XX  
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200216655-A2.  
XX  
PD 28-FEB-2002.  
XX  
PF 24-AUG-2001; 2001WO-US026585.  
XX  
PR 24-AUG-2000; 2000US-0227865P.  
PR 26-JAN-2001; 2001US-0264647P.  
PR 22-JUN-2001; 2001US-0300111P.  
XX  
PA (SRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
PI Harper JF, Kreps J, Wang X, Zhu T;  
XX  
DR WPI; 2002-304127/34.  
XX  
PT Identifying a stress condition to which a plant cell has been exposed and  
producing plants with increased tolerance to these abiotic stresses.  
XX  
PS Claim 144; SEQ ID NO 737; 577bp + Sequence Listing; English.  
XX  
CC The invention relates to identifying a stress condition to which a plant  
cell has been exposed, comprising: (a) contacting nucleic acid  
representative of expressed polynucleotides in the plant cell with an



CC array or probes representative of the plant cell genome; and (b)  
 CC detecting a profile of expressed polynucleotides in the plant cell  
 CC characteristic of a stress response. The method is useful in the  
 CC production of transgenic plants, cells and seeds and in producing plants  
 CC with increased tolerance to abiotic stress. The present sequence is that  
 CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
 CC in methods of the invention. Note: The sequence data for this patent is  
 CC not represented in the printed specification but is based on sequence  
 CC information supplied to Derwent by the European Patent Office  
 XX  
 XX  
 SQ Sequence 1005 BP; 287 A; 205 C; 235 G; 278 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	1.17e-72	Length:	1005
Score:	698.00	Matches:	131
Percent Similarity:	68.08%	Conservative:	46
Best Local Similarity:	50.38%	Mismatches:	75
Query Match:	52.13%	Indels:	8
DB:	6	Gaps:	1

US-10-624-061-16 (1-261) x ABZ12932 (1-1005)

QY	10	AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGluAsp	29
DB	223	GATGAGAGTGAGAGTTGACTCTTCAGAGTATTAGAACTCTTTGATCCGTCAGAGGAC	282
QY	30	ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp	49
DB	283	AGCATTATATTGGCTATTGGAGAGCCAGTACTGTTACAATGCTGATCTATGAT	342
QY	50	GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn	69
DB	343	CCTACTGCTTTGATGATGATGTTTCAATGCTCTTGGTGGTACATGTTGTTAAAGGC	402
QY	70	ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe	89
DB	403	ACTGAGAGCTTCACGCTAAGGTTGGTAGTTTAAAGTCTGATGAACATCTCTTCTC	462
QY	90	ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro	109
DB	463	CCTGATGATCACCAGAGCCATGTTGCTCTCTTCAGTACCACAAAGGTGTTCATTTT	522
QY	110	GlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu	129
DB	523	GCTGCTGATTCGATAAACAACATGAAGAGATGGAACATGACTTTCAGAGACCTGTT	582
QY	130	ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu	149
DB	593	CCAAGACTTGTGAAGAAAGCGGATGATGTTAACTACCGCTCAACAGCTGTCTGTGACGT	642
QY	150	SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys	169
DB	643	ATTCGCTTCACTGCTCTCTCAAGAGATTCATTCACGTTAAATTTGTCAGAGCTAA	702
QY	170	PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGlyLeu	189
DB	703	TTTCAAGCTCACCCGAGCATACGATCGCCCATCAAGCACAAGATAGGATGACATG	762
QY	190	MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys	209
DB	763	ATGATATGCTGATTCCTCCGACTGTGGAAGATGCGATAAAGAGAGAGTTGATGAA	822
QY	210	AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArg	229
DB	823	ACCCGAACATACGGGCAAGAGTGAAGTTGGATGGAGGAGAGAGAGAGAGAGAA	882
QY	230	-----LysPheAspProSerValAlaSerSerLeuTyrLys	241
DB	883	GAAGGAATGAATCTCATGTTTACAAATCAGTCCCATCTTATGCTGACTTATATGA	942
QY	242	AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp	261
DB	943	GATTGGATCATGCCCTTTAAACAAAGAGGTTCAAGTGGAGTACTTGTCTCAGAGACTGGAC	1002

## RESULT 9

ABZ1393	1005 BP; 287 A; 205 C; 235 G; 278 T; 0 U; 0 Other;
ID	ABA91393 standard; cDNA; 1207 BP.
XX	AC ABA91393;
XX	08-APR-2002 (first entry)
XX	Arabidopsis chorismate mutase-1 cDNA.
XX	Chorismate mutase-1; herbicide; plant; enzyme; gene; ss.
XX	Arabidopsis thaliana.
XX	WO200202798-A2.
XX	10-JAN-2002.
XX	22-JUN-2001; 2001WO-US020104.
XX	05-JUL-2000; 2000US-00610040.
XX	(PARA-) PARADIGM GENETICS INC.
XX	Boyes DC, Davis KR, Woessner JP, Gorlach J, Hamilton CM; Hoffmann NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R; Kjemtrup S;
XX	WPI; 2002-154754/20.
XX	Identifying compounds useful for modulating and inhibiting plant growth and development, by using chorismate synthase and chorismate mutase as targets which are essential for plant growth.
XX	Claim 37; Page 64; 73pp; English.
XX	The present sequence is that of a full-length cDNA encoding Arabidopsis thaliana chorismate mutase-1 (CM), an enzyme that catalyses the conversion of chorismate to prephenate in a branch of the shikimate pathway leading to phenylalanine and tyrosine biosynthesis. CM is essential for plant growth; inhibition of CM gene expression in plant seedlings results in severe chlorosis, reduced growth and developmental abnormalities. The invention provides compositions for the modulation of plant growth or development using CM antisense (see ABA91393) and sense polynucleotides, double-stranded RNA and ribozymes, and related expression cassettes and vectors. The compositions are particularly useful for the modulation and inhibition of plant growth, and for generating male sterile plants or plants that produce seedless fruit. CM is also a target for the identification of herbicides, and methods are provided for using CM in the identification of herbicide candidates. Methods are also provided for identifying compounds that stimulate the expression or function of CM. These can be used to promote growth and development
XX	Sequence 1207 BP; 343 A; 232 C; 278 G; 354 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.61e-72	Length:	1207
Score:	696.00	Matches:	131
Percent Similarity:	68.08%	Conservative:	46
Best Local Similarity:	50.38%	Mismatches:	75
Query Match:	51.98%	Indels:	8
DB:	6	Gaps:	1

US-10-624-061-16 (1-261) x ABA91393 (1-1207)

QY	10	AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGluAsp	29
DB	235	GATGAGAGTGAGAGTTGACTCTTGAAGGTATTAGAACTCTTTGATCCGTCAGAGGAC	294
QY	30	ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp	49



Db 295 AGCATTATTTGGCTATTGGAGAGAGCCAAAGTACTGTTACAACTGCTGATCTTATGAT 354  
 Qy 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValLysAsn 69  
 Db 355 CCTACTGCTTTGACATGGATGGTTCAATGGTTCTTTGGTTGAGTACATGGTTAAAGGC 414  
 Qy 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
 Db 415 ACTGAGAAGCTTCAGCTTAAGGTTGGTAGGTTTAAGAGTCTGATGAACATCTTCTTTC 474  
 Qy 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 Db 475 CTTGATGATCTACAGAGCTTATGCTCTCTCAAGAGAAATCCATTACGGTAAATTTGTTGAGAAGCTAAA 714  
 Qy 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrPheLysGluLeuLeu 129  
 Db 535 GCTGCTGATTTCGATAAACAATAACAAGAAGATATGGAACATGATACITTCAGAGACCTTGTT 594  
 Qy 130 ProLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 Db 595 CCAAGACTTGTGAAGAAGGCGATGATGGTAACCTACCGCTCAACAGCTGTCTGTGACGCT 654  
 Qy 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
 Db 655 ATCTGCTTCAAGTCTCTCAAGAGAAATCCATTACGGTAAATTTGTTGAGAAGCTAAA 714  
 Qy 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
 Db 715 TTTCAGAGCTCACCAGAGCATACAGTCCGCCATCAAGCAAGATCAAGATCAAGCTG 774  
 Qy 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
 Db 775 ATGGATATGCTGACATTCCTCCGACTGTGGAAGATCGGATGAAGAAGAGAGTGTGATGAAA 834  
 Qy 210 AlaValValPheGlyGlnGluValAlaAsnLeuAsnSerAspAsnAspAsnGluAsnArg 229  
 Db 835 ACCCGAATACAGGCGCAAGAGTGAAGTTGGAGTGGAGAGAGAAAGAAAGAAAGAA 894  
 Qy 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241  
 Db 895 GAAGGGAATGAATCTCATGTTTACAAAATCAGTCCGATCTTAGTTCGTGACITATATGA 954  
 Qy 242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
 Db 955 GATTGGATCATGCTTTTAAACAAAAGAGGTTCAAGTGGAGTACTTGTCTCAAGAGACTGGAC 1014  
 RESULT 10  
 ID ABA91399/c  
 XX ABA91399;  
 AC ABA91399;  
 XX 08-APR-2002 (first entry)  
 XX Arabidopsis chorismate mutase-1 antisense polynucleotide.  
 DE Arabidopsis chorismate mutase-1; enzyme; plant; antisense, ss.  
 XX Chorismate mutase-1; enzyme; plant; antisense, ss.  
 OS Arabidopsis thaliana.  
 XX WO200202798-A2.  
 XX 10-JAN-2002.  
 XX 22-JUN-2001; 2001WO-US020104.  
 XX 05-JUL-2000; 2000US-00610040.  
 XX (PARA-) PARADIGM GENETICS INC.  
 XX Boyes DC, Davis KR, Woessner JP, Goriach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 PI Kjemtrup S;

XX WPI; 2002-154754/20.  
 DR Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.  
 XX Claim 35; Page 71; 73pp; English.  
 XX The present sequence is that of an antisense oligonucleotide that is  
 CC specific for Arabidopsis thaliana chorismate-mutase-1 (CM) mRNA. CM  
 CC catalyzes the conversion of chorismate to prephenate in a branch of the  
 CC shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 CC is essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense and sense polynucleotides,  
 CC double-stranded RNA and ribozymes, and related expression cassettes and  
 CC vectors. The compositions are particularly useful for the modulation and  
 CC inhibition of plant growth, and for generating male sterile plants or  
 CC plants that produce seedless fruit  
 XX SQ Sequence 1207 BP; 354 A; 278 C; 232 G; 343 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,61e-72 Length: 1207  
 Score: 696.00 Matches: 131  
 Percent Similarity: 68.08% Conservative: 46  
 Best Local Similarity: 50.38% Mismatches: 75  
 Query Match: 51.98% Indels: 8  
 DB: Gaps: 1  
 US-10-624-061-16 (1-261) x ABA91399 (1-1207)

Qy 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 Db 973 GATGAGAGTGCAGAGTTTGACTCTTGAAGGTATTATGAAGACTCTTTGATCCGTCAAGAGGAC 914  
 Qy 30 ThrIleIleTyrGlyLeuIleGluAtgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 Db 913 AGCATATATTTGGCTATTGGAGAGAGCCAGTACTGTTACATGCTGATCTTATGAT 854  
 Qy 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValLysAsn 69  
 Db 853 CCTACTGCTTTGACATGGATGGTTTCAATGGTTCTTGTGTTGAGTACATGTTAAGGC 794  
 Qy 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
 Db 793 ACTGAGAAGCTTCAGCTAAGTTGGTAGGTTTGAAGTCTCTGATGAGACATCTTCTTCTTC 734  
 Qy 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 Db 733 CTTGATCATCTACAGAGCTTATGTTGCTCTCTTTCAGTACCCAAAGGTGTTCATTTT 674  
 Qy 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
 Db 673 GCTGCTGATTCGATTAACATAAACAAGAGATATGGAACATGATCTTCAGAGACCTTGTT 614  
 Qy 130 ProLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 Db 613 CCAAGACTTGTGAAGAAGGCGATGATGGTTAACTACGCTCAACAGCTGTCTGTGACGCT 554  
 Qy 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
 Db 553 ATCTGCTTCAAGTCTCTCAAGAGAAATCCATTACGGTAAATTTGTTGAGAAGCTAAA 494  
 Qy 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
 Db 493 TTTCAAGCTCTCACCAGAGCATACAGTCCGCCATCAAGCAAGATCAAGATCAAGCTG 434  
 Qy 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
 Db 433 ATGGATATGCTGACATTCCTCCGACTGTGGAAGTTCGATGAAGAAGAGAGTGTGATGAAA 374



QY	260	Leuasp	261	PR	18-JUN-1999;	99US-0139455P.
				PR	18-JUN-1999;	99US-0139456P.
DB	1018	TTGGAT	1023	PR	18-JUN-1999;	99US-0139457P.
				PR	18-JUN-1999;	99US-0139458P.
				PR	18-JUN-1999;	99US-0139459P.
RESULT 12				PR	18-JUN-1999;	99US-0139460P.
AAC33324				PR	18-JUN-1999;	99US-0139461P.
ID	AAC33324	standard; DNA; 1274 BP.		PR	18-JUN-1999;	99US-0139462P.
XX	XX			PR	18-JUN-1999;	99US-0139463P.
AC	AAC33324;			PR	18-JUN-1999;	99US-0139750P.
XX	XX			PR	18-JUN-1999;	99US-0139763P.
DT	17-OCT-2000	(first entry)		PR	21-JUN-1999;	99US-0139817P.
XX	XX			PR	22-JUN-1999;	99US-0139899P.
DE	Arabidopsis thaliana	DNA fragment SEQ ID NO: 2614.		PR	23-JUN-1999;	99US-0140353P.
XX	XX			PR	23-JUN-1999;	99US-0140354P.
KW	Hybridisation assay; genetic mapping; gene expression control;			PR	24-JUN-1999;	99US-0140895P.
KW	protein identification; signal transduction pathway; metabolic pathway;			PR	28-JUN-1999;	99US-0140823P.
KW	promoter; termination sequence; ss.			PR	29-JUN-1999;	99US-0140991P.
XX	XX			PR	30-JUN-1999;	99US-0141287P.
OS	Arabidopsis thaliana.			PR	01-JUL-1999;	99US-0141842P.
XX	XX			PR	01-JUL-1999;	99US-0142154P.
PN	EP1033405-A2.			PR	02-JUL-1999;	99US-0142055P.
PD	PD			PR	06-JUL-1999;	99US-0142390P.
XX	XX			PR	08-JUL-1999;	99US-0142803P.
XX	XX			PR	09-JUL-1999;	99US-0142920P.
XX	XX			PR	12-JUL-1999;	99US-0142977P.
XX	XX			PR	13-JUL-1999;	99US-0143542P.
XX	XX			PR	14-JUL-1999;	99US-0143624P.
XX	XX			PR	15-JUL-1999;	99US-0144005P.
XX	XX			PR	16-JUL-1999;	99US-0144085P.
XX	XX			PR	16-JUL-1999;	99US-0144086P.
XX	XX			PR	19-JUL-1999;	99US-0144325P.
XX	XX			PR	19-JUL-1999;	99US-0144331P.
XX	XX			PR	19-JUL-1999;	99US-0144332P.
XX	XX			PR	19-JUL-1999;	99US-0144333P.
XX	XX			PR	19-JUL-1999;	99US-0144334P.
XX	XX			PR	19-JUL-1999;	99US-0144335P.
XX	XX			PR	20-JUL-1999;	99US-0144352P.
XX	XX			PR	20-JUL-1999;	99US-0144632P.
XX	XX			PR	20-JUL-1999;	99US-0144884P.
XX	XX			PR	21-JUL-1999;	99US-0144814P.
XX	XX			PR	21-JUL-1999;	99US-0145086P.
XX	XX			PR	21-JUL-1999;	99US-0145088P.
XX	XX			PR	22-JUL-1999;	99US-0145085P.
XX	XX			PR	22-JUL-1999;	99US-0145087P.
XX	XX			PR	22-JUL-1999;	99US-0145089P.
XX	XX			PR	22-JUL-1999;	99US-0145192P.
XX	XX			PR	23-JUL-1999;	99US-0145145P.
XX	XX			PR	23-JUL-1999;	99US-0145218P.
XX	XX			PR	23-JUL-1999;	99US-0145224P.
XX	XX			PR	26-JUL-1999;	99US-0145276P.
XX	XX			PR	27-JUL-1999;	99US-0145913P.
XX	XX			PR	27-JUL-1999;	99US-0145918P.
XX	XX			PR	28-JUL-1999;	99US-0145919P.
XX	XX			PR	28-JUL-1999;	99US-0145951P.
XX	XX			PR	02-AUG-1999;	99US-0146386P.
XX	XX			PR	02-AUG-1999;	99US-0146388P.
XX	XX			PR	02-AUG-1999;	99US-0146389P.
XX	XX			PR	03-AUG-1999;	99US-0147038P.
XX	XX			PR	04-AUG-1999;	99US-0147204P.
XX	XX			PR	05-AUG-1999;	99US-0147302P.
XX	XX			PR	05-AUG-1999;	99US-0147192P.
XX	XX			PR	05-AUG-1999;	99US-0147260P.
XX	XX			PR	06-AUG-1999;	99US-0147303P.
XX	XX			PR	06-AUG-1999;	99US-0147416P.
XX	XX			PR	09-AUG-1999;	99US-0147493P.
XX	XX			PR	09-AUG-1999;	99US-0147935P.
XX	XX			PR	10-AUG-1999;	99US-0148171P.
XX	XX			PR	11-AUG-1999;	99US-0148319P.
XX	XX			PR	12-AUG-1999;	99US-0148341P.
XX	XX			PR	13-AUG-1999;	99US-0148565P.
XX	XX			PR	13-AUG-1999;	99US-0148684P.
XX	XX			PR	16-AUG-1999;	99US-0149368P.

PR 17-AUG-1999;	99US-0149175P.	Db	309	GATGAGAGTGGAGTTGACTCTTTGAAGGTATTAGAACTCTTTGATCCGTCAAGAGGAC	368
PR 18-AUG-1999;	99US-0149426P.	Qy	30	ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp	49
PR 20-AUG-1999;	99US-0149722P.	Db	369	AGCATTATATTGGCTATTGGAGAGGCAAGTACTTTTACATGCTGATCATCTTATGAT	428
PR 20-AUG-1999;	99US-0149929P.	Qy	50	GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn	69
PR 23-AUG-1999;	99US-0149930P.	Db	429	CCTACTGCTTTTGCATGATGCTTTCAATGGTCTTTGTTGGTGGTACATGTTTAAAGC	488
PR 26-AUG-1999;	99US-0150566P.	Qy	70	ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe	89
PR 27-AUG-1999;	99US-0151065P.	Db	489	ACTGAGAAGCTTCACGCTAAGGTTGGTAGGTTTAAAGAGTCTGTGATGAACATCTTCTTC	548
PR 27-AUG-1999;	99US-0151066P.	Qy	90	ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro	109
PR 30-AUG-1999;	99US-0151080P.	Db	549	CCTGATGATCTACGAGAGCTATGTCCTCTCTTCAGTACCCAAAGGTGTGATTTT	608
PR 30-AUG-1999;	99US-0151303P.	Qy	110	GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu	129
PR 31-AUG-1999;	99US-0151438P.	Db	609	GCTGCTGATTCGATAAACAATAACAAGAGATATGAACATGTACTTTCAGAGACCTTGT	668
PR 01-SEP-1999;	99US-0151930P.	Qy	130	ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu	149
PR 07-SEP-1999;	99US-0152363P.	Db	669	CCAGAGCTTGTGAAGAGGCGATGATGTAATTAAGGCTCAACAGCTGTCTGTGACGCT	728
PR 10-SEP-1999;	99US-0153070P.	Qy	150	SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys	169
PR 13-SEP-1999;	99US-0153758P.	Db	729	ATCTGCTTTCAGTGTCTCTCAAGAGAGATCCATTACGGTAAATTTGTCAGAACCTAAA	788
PR 15-SEP-1999;	99US-0154018P.	Qy	170	PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluLysLeu	189
PR 16-SEP-1999;	99US-0154032P.	Db	789	TTTCAAGCCTCAGGCAAGATACGAGTCCGCTCAACAGATGATGATGATGATGATGATG	848
PR 20-SEP-1999;	99US-0154779P.	Qy	190	MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys	209
PR 22-SEP-1999;	99US-0155139P.	Db	849	ATGGATATGCTGACATTCGCGACTGGAAGATGCGATAAAGAGAGATTTGAGATGAA	908
PR 23-SEP-1999;	99US-0155486P.	Qy	210	AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAspAsnArg	229
PR 24-SEP-1999;	99US-0155659P.	Db	909	ACCCGACATACGGCAAGAGTGAAGTGGGATGGAGAGAGAGAGAGAGAGAGAGAA	968
PR 28-SEP-1999;	99US-0156458P.	Qy	230	-----LysPheAspProSerValAlaSerSerLeuTyrLys	241
PR 29-SEP-1999;	99US-0157117P.	Db	969	GAAAGGAATGATCTCATGTTTACAAATCATGTCGAT-CTTAGTTGCTTATATGGA	1027
PR 05-OCT-1999;	99US-0157753P.	Qy	242	AsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp	261
PR 06-OCT-1999;	99US-0157865P.	Db	1028	GATTGGATCATGCTTTTAAACAAAGAGGTTCAAGTGGAGTACTTGTCTCAGAAAGACTGGAC	1087
PR 07-OCT-1999;	99US-0158029P.	RESULT 13			
PR 08-OCT-1999;	99US-0158232P.	ID	ABX93056	standard; cDNA; 1223 BP.	
PR 12-OCT-1999;	99US-0158369P.	AC	ABX93056;		
PR 13-OCT-1999;	99US-0159293P.	DT	22-MAY-2003	(first entry)	
PR 13-OCT-1999;	99US-0159294P.	DE	Corn chorismate mutase cDNA.		
PR 13-OCT-1999;	99US-0159295P.	KW	Corn; plant; gene; ss; chorismate mutase; aromatic amino acid; diet;		
PR 14-OCT-1999;	99US-0159330P.	XX	biosynthesis; chorismate; anthranilate; tryptophan; prephenate; tyrosine;		
PR 14-OCT-1999;	99US-0159331P.	XX	phenylalanine; prephenate dehydrogenase; herbicide; food crop.		
PR 14-OCT-1999;	99US-0159637P.	OS	Zea mays.		
PR 14-OCT-1999;	99US-0159638P.	XX	Key		
PR 18-OCT-1999;	99US-0159584P.	FT	Location/Qualifiers		
PR 21-OCT-1999;	99US-0160741P.	FT	3..998		
PR 21-OCT-1999;	99US-0160767P.	FT	/*tag= a		
PR 21-OCT-1999;	99US-0160770P.	FT	/product= "Chorismate mutase"		
PR 21-OCT-1999;	99US-0160814P.	FT	/partial		
PR 21-OCT-1999;	99US-0160815P.	FT	/note= "No start codon shown"		
PR 21-OCT-1999;	99US-0160902P.	XX			
PR 22-OCT-1999;	99US-0160981P.	XX			
PR 22-OCT-1999;	99US-0160989P.	XX			
PR 25-OCT-1999;	99US-0161404P.	XX			
PR 25-OCT-1999;	99US-0161405P.	XX			
PR 25-OCT-1999;	99US-0161406P.	XX			
PR 26-OCT-1999;	99US-0161359P.	XX			
PR 26-OCT-1999;	99US-0161360P.	XX			
PR 26-OCT-1999;	99US-0161361P.	XX			
PR 28-OCT-1999;	99US-0161920P.	XX			
PR 28-OCT-1999;	99US-0161922P.	XX			
PR 28-OCT-1999;	99US-0161933P.	XX			
PR 29-OCT-1999;	99US-0162142P.	XX			

Alignment Scores:  
 Pred. No.: 4.38e-70  
 Score: 677.50  
 Percent Similarity: 68.46%  
 Best Local Similarity: 50.38%  
 Query Match: 50.60%  
 DB: 3  
 Length: 1274  
 Matches: 131  
 Conservative: 47  
 Mismatches: 74  
 Indels: 9  
 Gaps: 1

US-10-624-061-16 (1-261) x AAC33324 (1-1274)  
 Qy 10 AppSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29



CC The present sequence is that of a full-length cDNA encoding Arabidopsis  
 CC thaliana chorismate mutase-3 (CM), an enzyme that catalyses the  
 CC conversion of chorismate to prephenate in a branch of the shikimate  
 CC pathway leading to phenylalanine and tyrosine biosynthesis. CM is  
 CC essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense (see ABA91401) and sense  
 CC polynucleotides, double-stranded RNA and ribozymes, and related  
 CC expression cassettes and vectors. The compositions are particularly  
 CC useful for the modulation and inhibition of plant growth, and for  
 CC generating male sterile plants or plants that produce seedless fruit. CM  
 CC is also a target for the identification of herbicides, and methods are  
 CC provided for using CM in the identification of herbicide candidates.  
 CC Methods are also provided for identifying compounds that stimulate the  
 CC expression or function of CM. These can be used to promote growth and  
 CC development  
 XX  
 SQ Sequence 1217 BP; 349 A; 273 C; 235 G; 360 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,18e-67 Length: 1217  
 Score: 654.50 Matches: 128  
 Percent Similarity: 67.98% Conservative: 44  
 Best Local Similarity: 50.59% Mismatches: 80  
 Query Match: 48.88% Indels: 1  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABA91395 (1-1217)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 DB 332 GATGAGTGGTACTTTGAACTTGAACCTGAAGCATTAGACACTCTTTGATTCGTCAAGAGGAC 391  
 QY 30 ThrIleIleTyrGlyLeuIleGlnArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 DB 392 AGTATTAATCTTTAACTCTTCTTGAACGAGCTCATGTCCTCAACGCTGATATATGAC 451  
 QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValValAsn 69  
 DB 452 GAGGATGCTTTACTATGGAAGGTTTCAGGATCTTTAGTTAGTTATGTTATGTCAGAA 511  
 QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAlaPhePhe 89  
 DB 512 ACTGAAAGCTTTCAGCGAAGGTGACAGGTACAGAGTCTGATGACATCCCTTTTC 571  
 QY 90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 DB 572 CCACATGCTTGCCTGACCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 631  
 QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTrpLysMetTyrPheLysGluLeuLeu 129  
 DB 632 TCGCGCGAATCGATAAATCAATCAACAAAGAGGTGGAATATGTTATTTCAACACACTTCTC 691  
 QY 130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 DB 692 CCCAGACTGGTCAACCGAGGGATGACCGTAATGTGTGCTCAGCTGCTCTCTGTGACACA 751  
 QY 150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169  
 DB 752 ATGTGTTTGCAGATACCTTTCAAGAGAGATTCATTTGCGTAAATTTGTTGCTGACGCCAAG 811  
 QY 170 PheArgAspAlaProGluAspTyrGluProIleAlaArgAlaLysAspLysGluGlyLeu 189  
 DB 812 TTTCGTGAAATCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871  
 QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValIleLysLys 209  
 DB 872 ATGCAACTTCTAACCTGCAAGACGCTTCAAGAGCTAGTCAAGAGAGAGAGTTGACATCAA 931  
 QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAsp---AsnGluAsn 228  
 DB 932 GCCAGATTTTGGTCAAGACATACATACGATTAACCGACCCAGAACTGAAGCTGATCTCTCC 991

QY 229 ArgLysPheAspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr 248  
 DB 992 TACAAATACACCTTAGCTTAGTTGCAAACTCTATGGAGAAAGGATCATGCCCTCACA 1051  
 QY 249 LysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
 DB 1052 AAGGAAGTCAAAATTGAGTACTTGTCTAGAGAGACTGGAT 1090  
 RESULT 15  
 ID ABA91401/c  
 ID ABA91401 standard; DNA; 1217 BP.  
 AC ABA91401;  
 DT 08-APR-2002 (first entry)  
 DE Arabidopsis chorismate mutase-3 antisense polynucleotide.  
 KW Chorismate mutase-3; enzyme; plant; antisense; ss.  
 OS Arabidopsis thaliana.  
 PN WO200202798-A2.  
 PD 10-JAN-2002.  
 PF 22-JUN-2001; 2001WO-US020104.  
 PR 05-JUL-2000; 2000US-00610040.  
 PA (PARA-) PARADIGM GENETICS INC.  
 PI Boyes DC, Davis KB, Woessner JP, Gorlach J, Hamilton CM;  
 PI Hoffman NE, Kloti AS, Zayed A, Ascenzi RA, Allen K, Mulpuri R;  
 PI Kjemtrup S;  
 DR WPI; 2002-154754/20.  
 XX Identifying compounds useful for modulating and inhibiting plant growth  
 PT and development, by using chorismate synthase and chorismate mutase as  
 PT targets which are essential for plant growth.  
 PS Claim 35; Page 72; 73pp; English.  
 CC The present sequence is that of an antisense oligonucleotide that is  
 CC specific for Arabidopsis thaliana chorismate-mutase-2 (CM) mRNA. CM  
 CC catalyses the conversion of chorismate to prephenate in a branch of the  
 CC shikimate pathway leading to phenylalanine and tyrosine biosynthesis. It  
 CC is essential for plant growth; inhibition of CM gene expression in plant  
 CC seedlings results in severe chlorosis, reduced growth and developmental  
 CC abnormalities. The invention provides compositions for the modulation of  
 CC plant growth or development using CM antisense and sense polynucleotides,  
 CC double-stranded RNA and ribozymes, and related expression cassettes and  
 CC vectors. The compositions are particularly useful for the modulation and  
 CC inhibition of plant growth, and for generating male sterile plants or  
 CC plants that produce seedless fruit  
 XX  
 SQ Sequence 1217 BP; 360 A; 235 C; 273 G; 349 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,18e-67 Length: 1217  
 Score: 654.50 Matches: 128  
 Percent Similarity: 67.98% Conservative: 44  
 Best Local Similarity: 50.59% Mismatches: 80  
 Query Match: 48.88% Indels: 1  
 DB: 6 Gaps: 1

US-10-624-061-16 (1-261) x ABA91401 (1-1217)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 DB 886 GATGAGTGGTACTTTGAACTTGAAGCATTAGACACTCTTTGATTCGTCAAGAGGAC 827







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: September 25, 2004, 02:59:07 ; Search time 90 seconds  
(without alignments)  
1609.358 Million cell updates/sec

Title: US-10-624-061-16  
Perfect score: 1339  
Sequence: 1 MAKAAEQSDSGNVYTLASV.....NWVPLTREVOVEVILLRLD 261

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues  
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=x1h  
-Q=/cgn2\_1/USPTO.spool/US10624061/runat.24092004.092407.16935/app.query.fasta\_1.455  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10624061 -CGN 1.1.69 @runat.24092004.092407.16935 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCUS COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1339	100.0	1020	4	US-09-454-279-15
2	818	61.1	1006	4	US-09-610-040-3
C	818	61.1	1006	4	US-09-610-040-9
	767.5	57.3	525	4	US-09-454-279-5
5	696	52.0	1207	4	US-09-610-040-2
C	696	52.0	1207	4	US-09-610-040-8
	685.5	51.2	1231	4	US-09-454-279-13
8	670.5	50.1	1223	4	US-09-454-279-11
9	654.5	48.9	1217	4	US-09-610-040-4
C	654.5	48.9	1217	4	US-09-610-040-10
	602	45.0	780	4	US-09-454-279-17
12	587	43.8	5176	4	US-09-610-040-6
					Sequence 15, Appl
					Sequence 3, Appl
					Sequence 9, Appl
					Sequence 5, Appl
					Sequence 2, Appl
					Sequence 8, Appl
					Sequence 13, Appl
					Sequence 11, Appl
					Sequence 4, Appl
					Sequence 10, Appl
					Sequence 17, Appl
					Sequence 6, Appl

13	549.5	41.0	5099	4	US-09-610-040-5	Sequence 5, Appli	
14	522	39.0	579	4	US-09-454-279-7	Sequence 7, Appli	
15	144.5	10.8	601	4	US-09-454-279-3	Sequence 3, Appli	
16	130	9.7	552	4	US-09-454-279-1	Sequence 1, Appli	
C	17	105	7.8	19250	4	US-08-961-527-35	Sequence 35, Appli
18	103	7.7	1830121	4	US-09-557-884-1	Sequence 1, Appli	
19	103	7.7	1830121	4	US-09-643-990A-1	Sequence 1, Appli	
20	96	7.2	1395	4	US-09-134-001C-1864	Sequence 1864, Ap	
21	96	7.2	2262	4	US-09-543-681A-850	Sequence 850, App	
22	92	6.9	3089	4	US-09-016-434-1330	Sequence 1330, Ap	
23	91.5	6.8	2394	3	US-09-319-989-7	Sequence 7, Appli	
24	90.5	6.8	2065	3	US-09-319-989-5	Sequence 5, Appli	
25	90.5	6.8	3081	3	US-09-319-989-9	Sequence 9, Appli	
26	90.5	6.8	3360	3	US-09-319-989-8	Sequence 8, Appli	
27	89.5	6.7	1773	4	US-09-107-532A-1028	Sequence 1028, Ap	
28	89	6.6	2427	4	US-09-543-681A-823	Sequence 823, App	
29	86.5	6.5	2793	4	US-09-134-001C-2477	Sequence 2477, Ap	
30	86.5	6.5	2976	4	US-09-386-962C-7	Sequence 7, Appli	
31	86	6.4	7953	4	US-08-956-171B-152	Sequence 152, App	
32	83.5	6.2	3414	1	US-07-973-320-1	Sequence 1, Appli	
33	83	6.2	2310	4	US-09-540-236-1190	Sequence 1190, Ap	
34	83	6.2	2709	4	US-09-107-532A-183	Sequence 183, App	
35	83	6.2	2749	1	US-08-344-695-1	Sequence 1, Appli	
36	83	6.2	3213	2	US-08-633-770A-4	Sequence 4, Appli	
37	83	6.2	3213	4	US-09-280-197-8	Sequence 8, Appli	
C	38	83	6.2	7425	4	US-09-453-702B-212	Sequence 212, App
C	39	83	6.2	119211	4	US-09-596-002-40	Sequence 40, Appli
40	82	6.1	3414	1	US-07-973-320-3	Sequence 3, Appli	
41	81.5	6.1	2286	4	US-09-112-580-1	Sequence 1, Appli	
42	81.5	6.1	580073	4	US-08-545-528D-1	Sequence 1, Appli	
43	81	6.0	1205	4	US-09-976-594-457	Sequence 457, App	
44	81	6.0	2559	4	US-09-071-035-325	Sequence 325, App	
45	81	6.0	2559	4	US-09-071-035-329	Sequence 329, App	

ALIGNMENTS

RESULT 1  
US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Lee Jian-Ming  
; APPLICANT: Famodu, Omolayo O.  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BBI299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Alignment Scores:	5.68e-163	Length:	1020
Pred. No.:	1339.00	Matches:	261
Score:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-10-624-061-16 (1-261) x US-09-454-279-15 (1-1020)

OY 1 MetalAllysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20

DB 23 ATGGCCAAAGCAGCAGACAAAGTCTTCTGGGAATGTGTACACGCTAGCTTCGTG 82

```
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArgAlaLys 40
Db 83 AGAGAGATTGGTTAGGCAAGAGATACCATTTATGGTCTCATTTAGAGAGCCCAAG 142
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60
Db 143 TTCCTCTAGCAATCTCACACCTATGATGAAGAAGTATGCTCAATCCAGGGTTTTGTGGC 202
QY 61 SerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
Db 203 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGCTGGAAGATAC 262
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100
Db 263 AAAAACCCCTGAAGAAACGCCCTTCTCCAGAAATTTACCAACATCAATTTGSCCATCT 322
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLysSerIle 120
Db 323 TACTCTCTCAACACAGTTTGTGATCTGCTGCTTCAATTAACATAAACAAGTCCATC 382
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140
Db 383 TGGAAATGTATTTCAAGAGTTACTTCCATTGCTTGCTACTTCCGGTGATGATGGCAAC 442
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160
Db 443 TATGGCAAACTGCAGCTAATGACCTTTCAATTTATGCACTCCATCTCTAGAAGGATTCAC 502
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
Db 503 TATGGAAGTTTGTAGCTGAGGTGAATTCAGGGATGCTCTCAAGACTACGAGCCCTTA 562
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200
Db 563 ATTCCAGCTAAGATAAAGAAGGATTTGATGAATTTGTTGACATTTACAAGCGTTGAAGAG 622
QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220
Db 623 ACGGTGAGGAAGAGAGTTGAAAGAAGGCTGTGGTGTGTTGGCGAGGAAGTGAATCTTAAC 682
QY 221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240
Db 683 AGTGATGACATGACAAATGAACCGTAAATTTGATCCATCAGTGGCTTTAGCTTGTAC 742
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeu 260
Db 743 AAAAAATGGGTGATACCTCTCACCAGGAGGTTTCAGGTGTGAGTACCTCTTGGCGCGCTA 802
QY 261 Asp 261
Db 803 GAC 805
```

```
RESULT 2
US-09-610-040-3
; Sequence 3, Application US/09610040
; Patent No. 6465217
; GENERAL INFORMATION:
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Goralach, Jörn
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Klori, Andreas
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS
; FILE REFERENCE: 9128.14
; CURRENT APPLICATION NUMBER: US/09/610.040
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 3
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-610-040-3
Alignment Scores:
Pred. No.: 8 81e-96 Length: 1006
Score: 818.00 Matches: 182
Percent Similarity: 72.83% Conservative: 31
Best Local Similarity: 61.13% Mismatches: 68
Query Match: 61.09% Indels: 4
DB: 4 Gaps: 2
US-10-624-061-16 (1-261) x US-09-610-040-3 (1-1006)
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18
Db 64 ATGGCAAGAGTCTTCGAATCGGATTCGGGTTCTGGTTGTTCCAAATGTACTGAGTCTGCAC 123
QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38
Db 124 TTAATCAGAGATCGTTGATAGGCAAGAACACCAATCGTCTTCAGCTGATTCAGAGA 183
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58
Db 184 GCTAAGTTTCCACTCAATCTCTCGTTTCGAGGAATCTCGTTGTCTAGATTCTGGAAGT 243
QY 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78
Db 244 TTCTCTCTCTCAGTGGTTCGTCAGAGAGACAGAAATCATCCAAGCTAAGGTAGGA 303
QY 79 ArgTyrLysAsnProGluGlnAlaPhePheProGluAsnLeuProProSerIleVal 98
Db 304 AGATATGAATACCGGAGAGAAATCTTTCTTCTTGAGAACATCTCTCACCTCGGTTTT 363
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLys 118
Db 364 CTTACGCAAAATATCCATCGGCTTTGCAACCTTAAGGCTCTATCTGTAAACATTAACAAA 423
QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138
Db 424 CAAATCTGGATATTTACTTTAAGAAATGCTTCTCTTTGTTGTCAAACCTGGCGATCAT 483
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158
Db 484 GGCAACTATCCATCAACTGCTCTAGTATCTCGGCTCTGTTTCAAGAGCTCTTTCGAGAGG 543
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGluAspTyrGlu 178
Db 544 ATTCACCTACGTAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCACAGATTACAG 603
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198
Db 604 CTTGGGATTCGGCTCAGGATAGAGAGGTTTGATGAAGCTGTTCACGCTTTCGAGAGATA 663
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218
Db 664 GAAGAAATGGTTAAGAGAGAGTGCAGAGAAAGAGAAACGTTTGGACAGAGATAAAA 723
QY 219 LeuAsnSerAspAsnAsnAsnGlnAsnArgLysPhe-----AspProSerValAla 236
Db 724 TTCAACTCTGGCTATGGCGATGAGAGTAAAGAAAGATATAAAGTGAATCCATTCGTTGCC 783
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256
Db 784 TCTGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCTT 843
QY 257 LeuArgArgLeuAsp 261
Db 844 CTACGTCGTCGAT 858
RESULT 3
```

US-09-610-040-9/c  
; Sequence 9, Application US/09610040  
; Patent No. 6465217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Goriach, Jörn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Kloti, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610, 040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 9  
; LENGTH: 1006  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-610-040-9  
  
Alignment Scores:  
Pred. No.: 8 81e-96 Length: 1006  
Score: 818.00 Matches: 162  
Percent Similarity: 72.83% Conservative: 31  
Best Local Similarity: 61.13% Mismatches: 68  
Query Match: 61.09% Indels: 4  
Gaps: 2  
DB: 4  
  
US-10-624-061-16 (1-261) x US-09-610-040-9 (1-1006)  
QY 1 MetAlaLysAlaAlaGluInSerProAspSerGly-----AsnValTyrThrLeuAla 18  
Db 943 ATGGCAAGACTCTTCGAATCGGATTCGGGTCTGGTGTTCCTCAATGCTACTGAGCTCTGAC 884  
  
QY 19 SerValArgGluAspLeuValArgGluAspThrIleIleTyrGlyLeuIleGluArg 38  
Db 883 TTAATCAGAAATCGTTGATAGCAGAGACACCATCGCTTCAGCTTGCATCGAGAGA 824  
  
QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
Db 823 GCTAAGTTTCCACTCAATCTCTGCTTCGAGGAATCTGTTGTCTAGATTCTGGAAGT 764  
  
QY 59 CysGlySerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
Db 763 TTCTCTTCTCTCACTGAGTTTTCGTCAGAGACAGAAATCATCAAGTAAGGTAGGA 704  
  
QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleVal 98  
Db 703 AGATATGAATACCGGAGAGAAATCCTTTCTCTTGAGAAATCTCTCACTCGGTTTTT 644  
  
QY 99 ProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLys 118  
Db 643 CCTACGCACAAATATCCATCGGCTTTTCACCCCTAAGGCTCTATCTGTAACTTAAACAA 584  
  
QY 119 SerIleTrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAsp 138  
Db 583 CAATCTGGGATATTTACTTTAAGAAATTCCTTCCTTTGTTGTCAACCTGGCGATGAT 524  
  
QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArg 158  
Db 523 GGCAACTATCCATCAACTGCTGCTAGTATCGCTGTTTACAAAGCTCTTTCCGAGAAGG 464  
  
QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
Db 463 ATTCACACGGAATAATTTAGCTAGGTCAAGTCAAGATGCTCCACAAAGATTACGAG 404  
  
QY 179 ProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerVal 198

Db 403 CCTCGGATTCGGCTCAGGATAGAGAGCTTTTGATGAAGCTTTTGAGCGTTTGAGAAAGTA 344  
  
QY 199 GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn 218  
Db 343 GAAGAAATGGTTTAAAGAGAGAGTSCAGAGAAAGACAGAAACGTTTGACAGAGATGAAA 284  
  
QY 219 LeuAsnSerAspAspAsnAsnGluAsnArgLysPhe-----AspProSerValAla 236  
Db 283 TTCACCTCTGGCTATGCGGATGAGAGTAAGAGAGAGTATAAAGTGGATCCATTCGTTGCC 224  
  
QY 237 SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
Db 223 TCTCGCATCTACGGGGAATGCTTATCCCTCTCACTAAGCTCGTGGTGGTGGATATCTT 164  
  
QY 257 LeuArgArgLeuAsp 261  
Db 163 CTACGTCGTCCTCGAT 149  
  
RESULT 4  
US-09-454-279-5  
; Sequence 5, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (432)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (514)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (516)  
US-09-454-279-5  
  
Alignment Scores:  
Pred. No.: 1 04e-89 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.29% Conservative: 0  
Best Local Similarity: 95.29% Mismatches: 7  
Query Match: 57.32% Indels: 4  
Gaps: 0  
DB: 4  
  
US-10-624-061-16 (1-261) x US-09-454-279-5 (1-525)  
QY 1 MetAlaLysAlaAlaGluInSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
Db 16 ATGGCCAAAGCAGACAGAAACAAAGTCTGATTCGGGAATGTGTACACGCTAGCTTCTGTG 75  
  
QY 21 ArgGluAspLeuValArgGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
Db 76 AGAGAGGATTTGGTTAGGCAAGAGATACCATCATTTATGGTCTCATTTGAGAGAGCCCAAG 135  
  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
Db 136 TTCCTTAGCAATCTCACACCTATGATGAAAGATGCTCAAAATCCAGGTTTTTTGTGGC 195  
  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80

Db 196 TCATTGGTGAATTTGTTTAAAGATACAGAGGCCATTCAGCTAAGCTTGAAGATAC 255  
 QY 81 LysAsnProGluGlnAlaPheProGluAsnLeuProSerIleValProSer 100  
 Db 256 AAAAAACCCCTGAAGAAACCCCTTCTCCAGAAAAATTTACCACCATCAATTTGGCCATCT 315  
 QY 101 TyrSerPheGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 Db 316 TACTCTCTCAACAGATTTTGGATCTCTGGTGGCTGCTTCAATACATTAACAGTCACT 375  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
 Db 376 GGGAAATGTATTTCAAGAGTTACTTCCATTTGCTTACTTCCGGTGATGATGNAAC 435  
 QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHi 160  
 Db 436 TAATGCGCAAACTGAGCTAATGACCT-TCATTATTGAGTCT-ATCTCTAGAGAT-CA 492

## RESULT 5

US-09-610-040-2

; Sequence 2, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woesner, Jeffrey

; APPLICANT: Gorlach, Jorn

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS

; FILE REFERENCE: 9128.14

; CURRENT FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1207

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-610-040-2

## Alignment Scores:

Pred. No.:	6,37e-80	Length:	1207
Score:	696.00	Matches:	131
Percent Similarity:	68.08%	Conservative:	46
Best Local Similarity:	50.38%	Mismatches:	75
Query Match:	51.98%	Indels:	8
DB:	4	Gaps:	1

US-10-624-061-16 (1-261) x US-09-610-040-2 (1-1207)

QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29  
 Db 235 GATGAGTGTGAGTGTGACTCTTGAAGGTATTAGAAACTCTTTGATCCGTCAGAGGAC 294  
 QY 30 ThrIleIleTyrGlyLeuLeuGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
 Db 295 AGCATATATTGGCTATTGGAGAGCCAGGACTGTTCATAAGTGTGATCTATTGAT 354  
 QY 50 GlutylsTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69  
 Db 355 CCTACTGCTTTTGACATGATGGTGGTTTCAATGGTCTTTGGTGGTGGTACATGTTAAAGGC 414  
 QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89

Db 415 ACTGAGAAGCTTCACCTAAGGTGGTAGTTTAAAGTCTCTGATGAACATCTTCTTC 474  
 QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
 Db 475 CCTGATGATCTACAGAGCTATGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 534  
 QY 110 GlyAlaAlaSerIleAsnLysSerIleTyrLeuLysSerIleTyrPheLysGlnLeuLeu 129  
 Db 535 GCTGCTGATTGATGAATCAACATTAACAGAGATATGAACATGATCTCTCAGACCTTGT 594  
 QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
 Db 595 CCAAGACTTGTGAAGAAAGGCGATGATGTAACACCGCTCAACAGCTGTCTGTGACGCT 654  
 QY 150 SerLeuLeuGlnSerIleSerArgArgIleHiSerTyrGlyLysPheValAlaGluValLys 169  
 Db 655 ATCTGCTTCTCAGTGTCTCTCAAGAGATCCATTACGGTAAATTTGTCAGAAGCTAAA 714  
 QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
 Db 715 TTTCAAGCTCACCAGAGCATACGAGTCGGCCATCAAGACACAGATAAGGATCGACTG 774  
 QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
 Db 775 ATGGATATGCTGACATTCCTCCGACTGTGGAAGATGCGATAAAGAGAGAGTTCAGATGAA 834  
 QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAsnGluAsnArg 229  
 Db 835 ACCCGAACATACGGCAAGAGTGAAGTTGGGATGGAGGAGAAAGAGAGAGAGAGAA 894  
 QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241  
 Db 895 GAAGGGAATGAATCTCATGTTTACAAATCATGTCGATCTTAGTTGGTGAATATATGGA 954  
 QY 242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
 Db 955 GATTGATCATGCTCTTAAACAAAGAGGTTCAAGTGGAGTCTTCTCTCAGAAGACTGGAC 1014

## RESULT 6

US-09-610-040-8/c

; Sequence 8, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woesner, Jeffrey

; APPLICANT: Gorlach, Jorn

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHASE

; FILE REFERENCE: 9128.14

; CURRENT FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; LENGTH: 1207

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-610-040-8

## Alignment Scores:

Pred. No.:	6,37e-80	Length:	1207
Score:	696.00	Matches:	131
Percent Similarity:	68.08%	Conservative:	46
Best Local Similarity:	50.38%	Mismatches:	75
Query Match:	51.98%	Indels:	8
DB:	4	Gaps:	1

US-10-624-061-16 (1-261) x US-09-610-040-8 (1-1207)

```
QY 10 AspSerGlyAsnValThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db 973 GATGAGTGTGAGTTGACTCTTGAAGGTATTAGAACTCTTGTGATCCGTCAAGAGGAC 914
QY 30 ThrIleIleTyrGlyLeuIleGluAlaGlyPheProSerAsnSerHisThrTyrAsp 49
Db 913 AGCATATATTGGCTATTGGAGAGACCAAGTACTGTTACAATGCTGATACCTATGAT 854
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValValAsn 69
Db 853 CCTACTGCTTTTGACATGATGCTTTCAATGCTCTTTGGTTGAGTACATGTTAAAGGC 794
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyTyrLysAsnProGluGluAsnAlaPhePhe 89
Db 793 ACTGAGAAGCTTCACGCTAAGTGTGAGTTTAAGAGTCTGATGAACATCCTTTCTTC 734
QY 90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db 733 CCTGATGATCTACAGAGCCTATGTTGCTCTCTTCTCAAGTACCCCAAGGTGTGATTTT 674
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGlnLeuLeu 129
Db 673 GCTGCTGATTCATTAACATTAACAAAGAGATATGGAACATGTAATTCAGAGACCTTGT 614
QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
Db 613 CCAAGACTTGTGAAGAAAGCGCATGATGTAATCACTACGGCTCAACAGCTGTCTGTGACGCT 554
QY 150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db 553 ATCTGCTTCAGTGTCTCTCAAGAGATCCATTAAGTAAATTTGTTGCAGAGCTTAA 494
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGlyLeu 189
Db 493 TTTCAAGCTCACCGAAGCATACGAGTCGCGCCATCAAGCAACAAGATGAAGATCGACTG 434
QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db 433 ATGGATATCTGACATCTCCGACTGCGAGATGCGATTAAGAGAGAGATTTGAGATGAA 374
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArg 229
Db 373 ACCCGACATACGGCAAGAGTGAAGTTGGGTGGAGGAGAGAAAGAGAGAGAA 314
QY 230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db 313 GAAGGAATGAATCTCATGTTTACAAATCACTGTCGATCTTAGTTGGTGAATATGGA 254
QY 242 AsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
Db 253 GATTGGATATGCTTTAAACAAAGAGTTCAAGTGGAGTACTTCTCTCAGAGAGCTGGAC 194
```

RESULT 7

```
US-09-454-279-13
; Sequence 13, Application US/09454279
; Patent No. 662798
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Onolayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1231
```

```
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-454-279-13
```

```
Alignment Scores: 1.49e-78 Length: 1231
Pred. No.: 685.50 Matches: 135
Score: 68.70% Conservative: 45
Best Local Similarity: 51.53% Mismatches: 81
Query Match: 51.19% Indels: 1
DB: Gaps: 1
```

US-10-624-061-16 (1-261) x US-09-454-279-13 (1-1231)

```
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20
Db 238 GTACCAAGAGAGAGAGAAACACAGAGATAGATCAAGTGAATACTGACCTTGGCAACATT 237
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40
Db 298 AGAACCTCTTGTGTAGGCAAGAGACAGCATCATATTACGCTCTTAGAGAGACACAG 357
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60
Db 358 TTTTGTCTACAAATGCTGATATATATGATAAAATGCTTTCATGATGGATTTGATGGC 417
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
Db 418 TCCTTGTGTAATTCATGTTAGAGAAACGAAACACTACATCAACAGGTTGGAGATAC 477
QY 81 LysAsnProGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100
Db 478 AAGAGCCTCATGAGCACCCATCTTTCGAGAGATCTGCTGAACACCTGTGGCCACCT 537
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db 538 CTCAGATCCAAAGGTTTTCATCTCTTTCGATCTTATTAATATCAACAGGAGATT 597
QY 121 TrpLysMetTyrPheLysGluLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140
Db 598 TGGAAATGATTTTGTGATGAGCTTCTTCCAAAGATTAGTGAAGAGAGAGTGTGTAAT 657
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160
Db 658 TATGGATCCAGTGTCTTGTGACAGATCTGCTGCGGCGCTCTCCAAAGAAATTCAC 717
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180
Db 718 TATGTTAAGTTTGTGGCAGAGGCTTAAGTTTCAAGAGTCTCTGAAGCTTACATGCTGCG 777
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200
Db 778 ATAATAGCACAGGACTGCGATCAACTAATGCCTCTCATCTATGAAACGGTGGAGCGT 837
QY 201 ThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsnLeuAsn 220
Db 838 GCTATTGAACATAGGGTGAAGCTTAAGCTTAAGTCTTTGGACAGGAGGTGGATTAGGC 897
QY 221 SerAspAspAsnAspAsnGluAsnArg---LysPheAspProSerValAlaSerSerLeu 239
Db 898 GCTGAAGACACACGGCGCTCCACCAATGTCACAGATAAGGCCAGTTTGGTGGCTGAACGT 957
QY 240 TyrLysAsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArg 259
Db 958 TACAGCTACAGGATCATGCCGCTAAACCAAGAGGTTCAAGTAGCCTACTTCTGCTGAGGAGA 1017
QY 260 LeuAsp 261
Db 1018 TTGGAT 1023
```

RESULT 8

```
US-09-454-279-11
; Sequence 11, Application US/09454279
```



150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluVallys 169  
752 ATGTGTTTGGAGATACCTTCAAGAGAAATTCACCTGGCTGAATTTGTTGCTGAGCCCAAG 811  
170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
812 TTTCTGAAATCTTGCTGCTATGAAACAGCTATCAAAAGAACCAAGACCGACACAGCTG 871  
190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
872 ATGCAACTTCACTACGAAACCGTTGAAAGAGTAGTCAAGAGAGAGAGTTGAGATCAAA 931  
210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAsp---AsnGluAsn 228  
932 GCCAGAAATTTGGTCAAGACATACGATTAAAGCCCAAGAACTGAAGCTGATCTTCC 991  
229 ArgLysPheAspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr 248  
992 TACAAATACAACTAGCTTAGTTGAAACCTTATGAGAAAGGATCATGCCCTCACA 1051  
249 LysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
1052 AAGGAAGTCCAAATGAGTACTTGGTTAGAACACTGGAT 1090  
RESULT 10  
US-09-610-040-10/c  
; Sequence 10, Application US/09610040  
; Patent No. 645217  
; GENERAL INFORMATION:  
; APPLICANT: Boyes, Douglas  
; APPLICANT: Davis, Keith  
; APPLICANT: Woessner, Jeffrey  
; APPLICANT: Goriach, Jörn  
; APPLICANT: Hamilton, Carol  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Klotz, Andreas  
; APPLICANT: Zayed, Adel  
; APPLICANT: Ascenzi, Robert  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS  
; TITLE OF INVENTION: CHORISMATE MUTASE EXPRESSION OR ACTIVITY IN PLANTS  
; FILE REFERENCE: 9128.14  
; CURRENT APPLICATION NUMBER: US/09/610,040  
; CURRENT FILING DATE: 2000-07-05  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 10  
; LENGTH: 1217  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-610-040-10  
Alignment Scores:  
Pred. No.: 1.45e-74 Length: 1217  
Score: 654.50 Matches: 128  
Percent Similarity: 67.98% Conservative: 44  
Best Local Similarity: 50.59% Mismatches: 80  
Query Match: 48.88% Indels: 1  
DB: 4 Gaps: 1  
US-10-624-061-16 (1-261) x US-09-610-040-10 (1-1217)  
QY 10 AspSerGlyAsnValTyrThrLeuAlaSerValArgLeuValArgGlnGluAsp 29  
DB 886 GATGAGTGAAGTATTTGAACACTTGAAGCATTTAGACACTCTTTGATCGTCAAGAGGAC 827  
QY 30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49  
DB 826 AGTATTATCTTAATCTTCTGACGAGCTCAGTATCGCTACAAAGCTGATGATCTATGAC 767  
QY 50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValLysAsn 69  
DB 766 GAGGATGCCCTTACTATGGAAGGCTTTCAGAGGATCTTTAGTTGAGTTTATGGTCAGAGAA 707

QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89  
DB 706 ACTGAAAGCTTCACCGAAAGGTGGACAGGTACAAGAGTCTTGATGAGCATCCCTTTTC 647  
QY 90 ProGluAsnLeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109  
DB 646 CCACAATGCTTGCTGAACCTATCTTCTCTCAATATCAATACCCACAGGTTTGGCATCGT 587  
110 GlyAlaAlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129  
DB 586 TGGCCGCAATCGATAAATCAACATCAACAGAGGTGTGGAATATGTATTTCAACACCTTCTC 527  
130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149  
DB 526 CCCAGACTGCTCAAGCCAGGGGATGACGTAATTTGTGTTACGCTGCTCTCTGTGACACA 467  
150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluVallys 169  
DB 466 ATGTGTTTGGAGATACCTTTCAGAGAAATTCACCTGGCTAAATTTGTTGCTGAGCCCAAG 407  
170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189  
DB 406 TTTCTGAAATCTTGCTGCTATGAAACAGCTATCAAAAGAACCAAGACCGACACAGCTG 347  
190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209  
DB 346 ATGCAACTTCTTAACGTACGAAACGGTTGAAGAACTAGTCAAGAGAGAGAGTTGAGATCAAA 287  
210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAsp---AsnGluAsn 228  
DB 286 GCCAGAAATTTGGTCAAGACATACGATTAAAGCCCAAGAACTGAAGCTGATCTCTTCC 227  
229 ArgLysPheAspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr 248  
DB 226 TACAAATACAACTAGCTTAGTTGAAACCTTATGAGAAAGGATCATGCCCTCACA 167  
249 LysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
DB 166 AAGGAAGTCCAAATGAGTACTTGGTTAGAACACTGGAT 128  
RESULT 11  
US-09-454-279-17  
; Sequence 17, Application US/09454279  
; Patent No. 6627798  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omolayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: BB1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 780  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
US-09-454-279-17  
Alignment Scores:  
Pred. No.: 4.3e-68 Length: 780  
Score: 602.00 Matches: 118  
Percent Similarity: 71.08% Conservative: 27  
Best Local Similarity: 57.94% Mismatches: 53  
Query Match: 44.96% Indels: 6  
DB: 4 Gaps: 2  
US-10-624-061-16 (1-261) x US-09-454-279-17 (1-780)  
QY 64 GluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnPro 83



10 GAGTTCTTCTGTCGGAAGCCGAGGTCCTGCGACGCAAGGCTGGACACTATCAAAAGCCA 69  
 84 GluGluAsnAlaPhePheProGluAsnLeuProSerIleValProSerTyrSerPhe 103  
 70 GAAGATTGTCCTCTCTCAAGATCTTCACACACCTGCTTTCTTACCAAAAGGTGCG 129  
 104 LysGlnPheLeuHisProGlyAlaAlaSerIleAsnLeuSerIleTyrPheMet 123  
 130 CCAAGGTTTGCACCCCTTTGCTTCATGCTGCTGCTGATGATGATGATGATGATGATG 189  
 124 TyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGln 143  
 190 TATTTCAATGAATGCTCTACCAATATTCTCGGATGCGGATGATGCGCAACTATGCGCAA 249  
 144 ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLys 163  
 250 ACAGTTGCTAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 309  
 164 PheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArgAla 183  
 310 TATGTTGCTGAGGTGAAATTCAGAGACGCGCTCAAGATTATAGCCACCAATACGTGCT 369  
 184 LysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrValArg 203  
 370 AAGGACACTAATGCTCTGATGAATCTTACTAAGTTTCAAGTTTCAAGTTTCAAGTTTCA 429  
 204 LysArgValGluLysLysAlaValPheGlyGlnGluValAsnLeu----- 219  
 430 AAGAGAGTAGAAG 489  
 220 AsnSerAspAspAsnAsp-----AsnGluAsnArgLysPheAspProSerValAlaSer 237  
 490 GCGAAGCAAGATGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTCT 549  
 238 SerLeuTyrLysAsnTyrPheProLeuThrLysGluValGlnValGluTyrLeuLeu 257  
 550 AAGCTATATGATATGCTGGTAATGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 609  
 258 ArgArgLeuAsp 261  
 610 CCGGCTCTTGAC 621

## RESULT 12

US-09-610-040-6

; Sequence 6, Application US/09610040

; Patent No. 6465217

; GENERAL INFORMATION:

; APPLICANT: Boyes, Douglas

; APPLICANT: Davis, Keith

; APPLICANT: Woessner, Jeffrey

; APPLICANT: Gorlach, John

; APPLICANT: Hamilton, Carol

; APPLICANT: Hoffman, Neil

; APPLICANT: Kloti, Andreas

; APPLICANT: Zayed, Adel

; APPLICANT: Ascenzi, Robert

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF CHORISMATE SYNTHAS

; FILE REFERENCE: 9128.14

; CURRENT FILING DATE: 2000-07-05

; NUMBER OF SEQ ID NOS: 11

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 5176

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-610-040-6

Alignment Scores:

Pred. No.:

Score: 7.07e-65

Length: 5176

Matches: 162

Percent Similarity: 36.50% Conservative: 30  
 Best Local Similarity: 30.80% Mismatches: 69  
 Query Match: 43.84% Indels: 268  
 DB: 4 Gaps: 6  
 US-10-624-061-16 (1-261) x US-09-610-040-6 (1-5176)  
 QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 DB 2542 ATGCGAAGAGTCTTCCGATCCGATTCGGTTCGTTGTTCCAAATGTTACTAGTCTTGAC 2601  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArg 38  
 DB 2602 TTAATCAGAGAATCGTTGATTAGGCAAGAAGACACCATCGTCTTCAGCTTGATCGAGAGA 2661  
 QY 39 AlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 DB 2662 GCTAAGTTTCCACTCAATTCCTCGTTCGAGGAATCTCGTTGCTAGATTCTGGAAGT 2721  
 QY 59 CysGlySerLeuValGluPheValLysValLysAsnThrGluAlaIleGlnAlaLys----- 76  
 DB 2722 TTCTCTTCTCTCACTCAGTTCCTCGTCAGAGAGACAGAAATCATCCAGCTAAGTTTGC 2781  
 QY 76 ----- 76  
 DB 2782 TTCCCAATTTAAAAAAGTATCCTTTGCTAAAAATTAGATACAGAGATATCAATGCTCGT 2841  
 QY 76 ----- 76  
 DB 2842 TTGATTCGGTTTGGTATAGCATTCGTTTAGATTGTTCCATGAAATTAGCAGAGAAGTAAG 2901  
 QY 76 ----- 76  
 DB 2902 CTACAAGTCACTTCGATTCGAGTTTAAATAAGCCTGGATTCTTGAAATTAGCATGCTTTT 2961  
 QY 76 ----- 76  
 DB 2962 GTTTCGTATGTCCTCCTCCATTCGAAAGATGATACTTCGCTTGGCTGTATAATCTC 3021  
 QY 76 ----- 76  
 DB 3022 ATTGTGTGATAACTTCTTGTGTTTGTGAGTGGCAATCTGCCAATAAAAGGCTCCGACT 3081  
 QY 76 ----- 76  
 DB 3082 TTATCATATGATACAGAGATTTCCTTATGAAAACTCATATTATATGTGGAGATTGGAATG 3141  
 QY 76 ----- 76  
 DB 3142 GAGGACTATTGTTTCTATTATTATATGCTGAAAGTCITATTTCATTATAATATTTCAT 3201  
 QY 76 ----- 76  
 DB 3202 CTCATGTTTATATCTTAAAGTTTCTGGATATTGAGCCTATATGTTTGTTCATTTGTTT 3261  
 QY 77 ----- 76  
 DB 3262 ACTTGAAACCTTATGTTGTTGTTATATTATATAGTAGGAGAGATATGATATACCCGAA 3321  
 QY 85 GluAsnAlaPhePheProGluAsnLeuProProSerIleValProSerTyrSerPheLys 104  
 DB 3322 GAGAAATCCTTTCTTCTTCTGAGAACATCTCCTCAGTTCGTTTCTTCTAGCAGCAATATCCA 3381  
 QY 104 ----- 104  
 DB 3382 TCGGTATGATAGTAAAGTCTTGAGCAATTTTCTTAGAATCTGAAATGCTTAGTCTA 3441  
 QY 105 -----GlnPheLeuHisProGlyAlaAlaSerIleAsnIle 116  
 DB 3442 ACAGTACTTTTCTTCTTCTTGTAGTTAGGCT-TTGACCCCTAAGGCTCTATCTGTTAACTT 3500  
 QY 117 AsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGly 136







```

FEATURE:
; NAME/KEY: unsure
; LOCATION: (526)
FEATURE:
; NAME/KEY: unsure
; LOCATION: (547)
FEATURE:
; NAME/KEY: unsure
; LOCATION: (566)
FEATURE:
; NAME/KEY: unsure
; LOCATION: (579) .. (580)
FEATURE:
; NAME/KEY: unsure
; LOCATION: (582)
US-09-454-279-3

```

```

Alignment Scores:
Pred. No.: 2.93e-09 . Length: 601
Score: 144.50 Matches: 41
Percent Similarity: 59.43% Conservative: 22
Best Local Similarity: 38.68% Mismatches: 34
Query Match: 10.79% Indels: 10
DB: 4 Gaps: 2

```

US-10-624-061-16 (1-261) x US-09-454-279-3 (1-601)

```

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20
DB 231 GTGACCAAGGAGGAGAAACAGAGGATAGATCAAGTGAATACTGACCTTGGACACATT 290
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40
DB 291 AGAACCTCCTTGGTTAGGCAAGACAGCATCATATTGAGGCTCTTAGAGAGACACAG 350
QY 41 -PheProSerAsnSerHisThrTyrAspGluLys-TyrAlaGlnIleGlnGlyPheCysG 60
DB 351 TTTTGCTACAAATGCTGATATATATGATAAAATGCTTCCATGTGGATGGATTGATG 410
QY 60 LysSerLeuValGluPheValVal--LysAsnThrGluAlaIleGlnAlaLysAlaGlyA 79
DB 411 GCCTTTGGTTNAATTCATGGTTANGAGAAACCGAAACCTACATCANCAGTTTGGGA 470
QY 79 IGTyrlLysAsnProGluGluAsnAlaPhePheProGluAsnLeuPro----- 94
DB 471 GATACAGAGCCCTGATGANCACCA-TTCTTTCCGGANGATCTGCTGAACAATGTNGCA 529
QY 95 -----ProSerIle 97
DB 530 ACCTCCCCAGTATC 543

```

Search completed: September 25, 2004, 05:06:03  
Job time : 112 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 04:17:54 ; Search time 514 Seconds  
(without alignments)  
2571.891 Million cell updates/sec

Title: US-10-624-061-16

Perfect score: 1339

Sequence: 1 MAKAAEQPSDGNVTLASV.....NWVPLTRKQVQVYLLRDL 261

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10624061/runat\_24092004\_092409\_17014/app\_query.fasta\_1.455  
-DB=published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOPCPL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODES=LOCAL -OUTFMT=ptc -NCRM=ext -HEARSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -MAP=US10624061.cgn\_1\_1\_537\_runat\_24092004\_092409\_17014  
-NCFU=6 -ICPU=3 -NO USSAR -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/prodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1

US-09-454-279-15  
; Sequence 15, Application US/09454279  
; Publication No. US20020184658A1  
; GENERAL INFORMATION:  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Famodu, Omelayo O.  
; APPLICANT: Lee, Jian-Ming  
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
; FILE REFERENCE: B1299 US NA  
; CURRENT APPLICATION NUMBER: US/09/454,279  
; CURRENT FILING DATE: 1999-12-03  
; EARLIER APPLICATION NUMBER: 60/110,845  
; EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-454-279-15

Alignment Scores:

Pred. No.: 3 54e-164 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0

US-10-624-061-16 (1-261) x US-09-454-279-15 (1-1020)

QY 1 MetAlaLysAlaAlaGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 DB 23 ATGGCCAAAGCAGCAGAAAGTCTGATTCGGGAATGTGTACACGCTAGCTTCTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40  
 DB 83 AGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGGTCTCATTTGAGAGCCCAAG 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 DB 143 TTCCCTAGCAATTTCTCACCACTATGATGAAAAGTATGCTCAATCCAGGCTTTTGTGGC 202  
 QY 61 SerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 DB 203 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCAATTCAGCTTAAGCTGGAAGATAC 262  
 QY 81 LysAsnProGluGluAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 263 AAAAACCCCTGAAGAAACGCTTCTCCAGAAAATTTTACCACCATCAATTTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 DB 323 TACTCCTTCAACACGATTTTGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382  
 QY 121 TrpLysMetTyrPheLysGlnLeuProLysLeuValThrSerGlyAspGlyAsn 140  
 DB 383 TGGAAAATGTATTTCAAGAGTTACTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHis 160  
 DB 443 TATGGCAAACTGCAAGCTTCTTCCAGAAAATTTTACCACCATCAATTTGTGCCATCT 502  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 503 TATGGAAAATGTATTTCAAGAGTTACTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 562  
 QY 181 IleArgAlaLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 200  
 DB 623 ACGGTGAGAAAGAGAGTTGAAAAGAGGCTGTGGTGTGGTGGTGGTGGTGGTGGTGGT 682  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
 DB 683 AGTGATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGTAC 742  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGlnTyrLeuLeuArgArgLeu 260  
 DB 743 AAAAATGGGTGATACCTCTCCACCAAGAGGTTTCAGGTGAGTACCTCTTTCGCGCGTCTA 802  
 QY 261 Asp 261  
 DB 803 GAC 805

## RESULT 2

US-10-624-061-15  
 ; Sequence 15, Application US/10624061  
 ; Publication No. US20040019929A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Falco, S. Carl  
 ; APPLICANT: Famedu, Omolayo O.  
 ; APPLICANT: Lee, Jian-Wing

; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES  
 ; FILE REFERENCE: B51299 US NA  
 ; CURRENT APPLICATION NUMBER: US/10/624,061  
 ; CURRENT FILING DATE: 2003-07-21  
 ; PRIOR APPLICATION NUMBER: US/09/454,279  
 ; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: Microsoft Office 97  
 ; SEQ ID NO 15  
 ; LENGTH: 1020  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; US-10-624-061-15

## Alignment Scores:

Pred. No.: 3 54e-164 Length: 1020  
 Score: 1339.00 Matches: 261  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-10-624-061-16 (1-261) x US-10-624-061-15 (1-1020)

QY 1 MetAlaLysAlaAlaGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 DB 23 ATGGCCAAAGCAGCAGAAAGTCTGATTCGGGAATGTGTACACGCTAGCTTCTGTG 82  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLys 40  
 DB 83 AGAGAGGATTTGGTTAGGCAAGAGGATACCATTTATGGTCTCATTTGAGAGCCCAAG 142  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 DB 143 TTCCCTAGCAATTTCTCACCACTATGATGAAAAGTATGCTCAATCCAGGCTTTTGTGGC 202  
 QY 61 SerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 DB 203 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCAATTCAGCTTAAGCTGGAAGATAC 262  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 263 AAAAACCCCTGAAGAAACGCTTCTTCCAGAAAATTTTACCACCATCAATTTGTGCCATCT 322  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 DB 323 TACTCCTTCAACACGATTTTGGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 382  
 QY 121 TrpLysMetTyrPheLysGlnLeuProLysLeuValThrSerGlyAspGlyAsn 140  
 DB 383 TGGAAAATGTATTTCAAGAGTTACTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHis 160  
 DB 443 TATGGCAAACTGCAAGCTTCTTCCAGAAAATTTTACCACCATCAATTTGTGCCATCT 502  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 503 TATGGAAAATGTATTTCAAGAGTTACTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTG 562  
 QY 181 IleArgAlaLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 200  
 DB 563 ATTCGAGCTTAGGATAAAGAGGATTTGATGAAAATTTTGTGACATTTACAGGCTTGAAGAG 622  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 DB 623 ACGGTGAGAAAGAGAGTTGAAAAGAGGCTGTGGTGTGGTGGTGGTGGTGGTGGTGGT 682  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
 DB 683 AGTGATGACATGACATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGTAC 742

QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgGlu 260  
Db 743 AAAAATGGGTGATACCTCTCACCAAGAGGTTTCAGGTTGAGTACCTCTTGGCCGCTCTA 802  
QY 261 Asp 261  
Db 803 GAC 805  
RESULT 3  
US-10-425-114-7639  
; Sequence 7639, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 7639  
; LENGTH: 1079  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700677130\_FLI  
US-10-425-114-7639  
Alignment Scores:  
Pred. No.: 3,86e-164 Length: 1079  
Score: 1339.00 Matches: 261  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 13  
US-10-624-061-16 (1-261) x US-10-425-114-7639 (1-1079)  
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
Db 110 ATGCCAAAGACGAGCAACAAAGCTCTGATTCCTGGGAATGTGTACACGCTAGCTCTGTG 169  
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
Db 170 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGCTCTCATTTGAGAGAGCCAA 229  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
Db 230 TTCCCTAGCAATTCACACCTATGATGAAAGATGTGCTCAAAATCCAGGGTTTTGTGGC 289  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
Db 290 TCATTTGGTGAATTTGTTGATTAAGAAATACAGAGCCATTCAGCTAAGCTGGAAGATAC 349  
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProPheSerIleValProSer 100  
Db 350 AAAAACCTCGAAGAAACGCGCTCTCCAGAAATTTACCAACATCAATTTGGCCATCT 409  
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
Db 410 TACTCTTCAACAGTTTGTTCATCTGCTGCTGCTCAATTAACATAAACAAGTCCATC 469  
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140  
Db 470 TGGAAATGTATTTCAAGAGATTTACTTCCATTTGCTTACTTCCGGGTGATGATGGCAAC 529  
QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHis 160

Db 530 TATGCCAAACGCGAGCTAATGACCTTTCATTTATTCAGTCCATCTCTAGAAGGATTCAC 589  
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
Db 590 TATGAAAGTTTGTAGCTGAGGTGAATTCAGGAGTCTCTCAAGACTACGAGCCCTTTA 649  
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
Db 650 ATTCGAGCTAAGGATAAAGAGGATTCATGAAATTTGACATTTTACAAGCGTTGAAGAG 709  
QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
Db 710 ACGTGAGGAGAGAGTTGAAAGAGCGCTGTGCTGTTTGGCGAGGAGTGAATCTTAAC 769  
QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
Db 770 AGTGATGACAATGACAATGAAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTGTGAC 829  
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgGlu 260  
Db 830 AAAAATTTGGTGATACCTCTCACAAAGAGGTTTCAGTTGAGTACCTCTTGGCCGCTCTA 889  
QY 261 Asp 261  
Db 890 GAC 892  
RESULT 4  
US-10-424-599-93036  
; Sequence 93036, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 93036  
; LENGTH: 1653  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55025C.1  
US-10-424-599-93036  
Alignment Scores:  
Pred. No.: 1.01e-160 Length: 1653  
Score: 1315.00 Matches: 255  
Percent Similarity: 100.00% Conservative: 6  
Best Local Similarity: 97.70% Mismatches: 0  
Query Match: 98.21% Indels: 0  
Gaps: 13  
US-10-624-061-16 (1-261) x US-10-424-599-93036 (1-1653)  
QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
Db 447 ATGCCAAAGACGAGCAACAAAGCTCTGATTCCTGGGAATGTGTACACGCTAGCTCTGTG 506  
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
Db 507 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATTTATGCTCTCATTCAGAGAGCCAA 566  
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
Db 567 TTCCCTAGCAATTCACACCTATGATGAAAGATGTGCTCAAAATCCAGGGTTTTGTGGC 626  
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80

Db 627 TCATTGGTAGAATTGTTGTTAAGATAACAGAGGCCATTCAAGCTGAAGCTGGAAGATAC 686  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 Db 687 AAAAACCCTGAGAAAACCCCTCTTCCAGAAAATTTACCAACCAATCAATTGTGCCATCT 746  
 QY 101 TyrSerPheGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 Db 747 TACTCCTTCAAAACAGTTTTTGGCATCTGGAGCTGCTTCAATTAACATTAAACAGTCCATC 806  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
 Db 807 TGGAAATATATTCCAGAGGTACTTCCATTGCTTGTCTTACTTGGGGTGATGGAAC 866  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHis 160  
 Db 867 TATGCACAACCTGCAGCTAATCATCTTTCATTTTGCAGGCCATCTCTAGAGGATTCAC 926  
 QY 161 TyrGlyPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 Db 927 TATGGAAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGACTACGAGCCCTTTA 986  
 QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
 Db 987 ATTCCAGCTAAGATTAAGAGGATGATGAATTTGTGACATTTACAGCGTTGAAGAG 1046  
 QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220  
 Db 1047 ACGGTGAGAGAGGCTTGAAGAAGGCTGTGGTGTGTGGCGAGGAAGTCAATCTTGAC 1106  
 QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerIleTyr 240  
 Db 1107 AATGAGGAGATGATGAATGAACCGTAAATTTGATCCATCAGTGGCTTCTAGCTTGTAC 1166  
 QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeu 260  
 Db 1167 AAAAATGGGTGATACCCCTTACCAAGGAGGTTGAGTTGAGTACCTATTGGCGGCTTA 1226  
 QY 261 Asp 261  
 Db 1227 GAC 1229

## RESULT 5

US-10-425-114-15059  
 ; Sequence 15059, Application US/10425114  
 ; Publication No. US20040034888A1

## GENERAL INFORMATION:

; APPLICANT: Liu, Jingdong  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Screen, Steven E  
 ; APPLICANT: Tabaska, Jack E  
 ; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 15059

; LENGTH: 992

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3050-017-E9\_FLI

US-10-425-114-15059

## Alignment Scores:

Pred. No.: 4.51e-116 Length: 992  
 Score: 970.00 Matches: 187  
 Percent Similarity: 97.99% Conservative: 8  
 Best Local Similarity: 93.97% Mismatches: 4

Query Match: 72.44% Indels: 0  
 DB: 13 Gaps: 0

US-10-624-061-16 (1-261) x US-10-425-114-15059 (1-992)

QY 63 ValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsn 82  
 Db 2 GTAGATTGTTCTTAAGATACAGAGGCCATTCAAGCTAAGCTGGAAGATCAAAAC 61  
 QY 83 ProGluGlnAsnAlaPhePheProGluAsnLeuProProSerIleValProSerTyrSer 102  
 Db 62 CCTGAAGAAAACCCCTCTTCCAGAAAATTTACCAACCAATCAATTGTGGCATCTTACTCC 121  
 QY 103 PheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIleTrpLys 122  
 Db 122 TTCAAACAGTTTTTGGCATCTCTGGAGCTGCTTCAATTAACATTAAACAGTCCATCTGGAAA 181  
 QY 123 MetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsnTyrAla 142  
 Db 182 ATGTATTTCAGAGGTACTTCCATTGTTGCTTCTCGGGGATGATGGAACATATGCA 241  
 QY 143 GlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArgIleHisTyrGly 162  
 Db 242 CAAACTGCAAGCTAATGATCTTTCATTTTGCAGGCCATCTCTAGAGGATTCACATGGA 301  
 QY 163 LysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArg 182  
 Db 302 AAGTTTGTAGCTGAGTGAAATTCAGGATGCTCTCAAGACTACGAGCCTTTAAATCGA 361  
 QY 183 AlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrVal 202  
 Db 362 GCTAAGATTAAGAGGATGATGAATTTGTGACATTTACAAGCGTTTCAAGAGACGGTG 421  
 QY 203 ArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsnSerAsp 222  
 Db 422 AGGAAGAGGTTTGAAGAAGGCCCGCAGTGTTTGGCAGGAGGATGAGTCTTGCACATGAG 481  
 QY 223 AspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyrLysAsn 242  
 Db 482 GACGATGATGAGAAAACCATTAAGTTTGTATCCATCAGTGGCTTCTAGCTTGTACAAAAT 541  
 QY 243 TrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261  
 Db 542 TGGGTGATACCCCTTACCAAGGAGGTTTCAAGTTGAGTACCTATTGGCGGCTTAGAC 598

## RESULT 6

US-09-938-842A-972

; Sequence 972, Application US/09938842A

; Patent No. US20020160378A1

## GENERAL INFORMATION:

; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun  
 ; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; PRIOR FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 972

; LENGTH: 798

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-972

Alignment Scores:



Pred. No.: 2,14e-96 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 9 Gaps: 2

US-10-624-061-16 (1-261) x US-09-938-842A-972 (1-798)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 DB 1 ARGCAAGAGCTTCGAAATCGGATTCGGTTCCTGTTCCAAATGCTACGATCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 DB 61 TTAATCAGAGATCGTTGATTAGGCAAGACAGACCATCGTCTTCAGCTTGTACGAGAGA 120  
 QY 39 AlalysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 59  
 DB 121 GCTAAGTTTCACATCAATTCCTCGTTCGAGGAATCTCGTTGTCTGAGATTCGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 DB 181 TTCCTCTCTCCTCAGCTGTTTTCGTCAGAGACAGAAATCATCCAGCTTAGGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluAlaPhePheProGluAsnLeuProSerIleVal 98  
 DB 241 AGATATGAATACCCGGAAGAGAAATCCCTTCTTCCTCGAGAACATTCCTCACTCGGTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLysHisProGlyAlaAlaSerIleAsnLys 118  
 DB 301 CCTACGCCAATATCCATCGCTTTCACCCCTAAGCTCTAICTGTATACATTAACAA 360  
 QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 DB 361 CAAATCTCGGATATTTACTTTAAGAAATTCGTTCTTGTTCGAAACCTGCGCATGAT 420  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleVal 158  
 DB 421 GGCAACTATCCATCAATCGCTGCTAGTCAAAATTCAGAGATGCTCCACAAAGATTACGAG 540  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 DB 481 ATTCACATCGGTAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCAAAGATTACGAG 540  
 QY 237 SerSerLeuTyrLysAsnTyrValIleProLeuThrLysGluValGlnValGluTyrLeu 256  
 DB 721 TCTCGCATCTACGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTGAGTATCTT 780  
 QY 257 LeuArgArgLeuAsp 261  
 DB 781 CTACGTCGCTCGAT 795

# RESULT 7

US-09-938-842A-972  
 ; Sequence 972, Application US/09938842A  
 ; Publication No. US2004000947629  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harper, Jeff  
 ; APPLICANT: Kreps, Joel  
 ; APPLICANT: Wang, Xun

APPLICANT: Zhu, Tong  
 TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
 FILE REFERENCE: SAME, AND METHODS OF USE  
 CURRENT APPLICATION NUMBER: US/09/938,842A  
 CURRENT FILING DATE: 2001-08-24  
 PRIOR APPLICATION NUMBER: US 60/227,866  
 PRIOR FILING DATE: 2000-08-24  
 PRIOR APPLICATION NUMBER: US 60/264,647  
 PRIOR FILING DATE: 2001-01-16  
 PRIOR APPLICATION NUMBER: US 60/300,111  
 PRIOR FILING DATE: 2001-06-22  
 NUMBER OF SEQ ID NOS: 5379  
 SEQ ID NO 972  
 LENGTH: 798  
 TYPE: DNA  
 ORGANISM: Arabidopsis thaliana  
 US-09-938-842A-972

Alignment Scores:  
 Pred. No.: 2,14e-96 Length: 798  
 Score: 818.00 Matches: 162  
 Percent Similarity: 72.83% Conservative: 31  
 Best Local Similarity: 61.13% Mismatches: 68  
 Query Match: 61.09% Indels: 4  
 DB: 11 Gaps: 2

US-10-624-061-16 (1-261) x US-09-938-842A-972 (1-798)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGly-----AsnValTyrThrLeuAla 18  
 DB 1 ATGCNAGAGCTTCGAAATCGGATTCGGTTCCTGTTGTTCCAAATGCTACGATCTTGAC 60  
 QY 19 SerValArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArg 38  
 DB 61 TTAATCAGAGATCGTTGATTAGGCAAGAGACACCATCGTCTTCAGCTTGTATCGAGAGA 120  
 QY 39 AlalysPheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPhe 58  
 DB 121 GCTAAGTTTCACATCAATTCCTCGTTCGAGAAATCTCGTTGTCTAGATTCGGAAGT 180  
 QY 59 CysGlySerLeuValGluPheValLysAsnThrGluAlaIleGlnAlaLysAlaGly 78  
 DB 181 TTCCTCTCTCCTCAGTATTTTCGTCAGAGACAGAAATCATCCAAAGTTAGGTAGGA 240  
 QY 79 ArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeuProSerIleVal 98  
 DB 241 AGATATGAATACCCGGAAGAGAAATTCCTTCTTCTTGAGAACATTCCTCACTCGGTTT 300  
 QY 99 ProSerTyrSerPheLysGlnPheLysHisProGlyAlaAlaSerIleAsnLys 118  
 DB 301 CCTACGCCAATATCCATCGCTTTCACCCCTAAGCTCTAICTGTATACATTAACAA 360  
 QY 119 SerIleTyrLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAsp 138  
 DB 361 CAAATCTCGGATATTTACTTTAAGAAATTCGTTCTTGTTCGAAACCTGCGCATGAT 420  
 QY 139 GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg 158  
 DB 421 GGCAACTATCCATCAATCGCTGCTAGTCAAAATTCAGAGATGCTCCAAAGATTACGAG 480  
 QY 159 IleHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGlu 178  
 DB 481 ATTCACATCGGTAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCAAAGATTACGAG 540  
 QY 179 ProLeuIleArgAlaLysAspLysGluLysLeuMetLysLeuLeuThrPheThrSerVal 198  
 DB 541 CCTCGCATCTACGGGTACGAGGCTTTCATGAAGCTGTTGACGTTTGAGAAAGTA 600  
 QY 199 GluGluThrValArgLysArgValGluLysLysAlaValPheGlyGlnGluValAsn 218  
 DB 601 GAAGAAATGGTTAAGAAAGAGAGATGCGCAAGAAAGCAAGAAACGTTTGGACAAGTAATA 660

Qy	219	LeuAenSerAaSpaAaSpaAaSpaAaArgLysPhe-----AapProSerValAla	236
Db	661	TTCACTCTGGCTATGGCCATCAGAGTAAGAAGAAGATATAAGTGGATCCATTGCTTGC	720
Qy	237	SerSerLeuTyLysAaSnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu	256
Db	721	TTCTGCATCATCGGGGAATGGCTTATCCCTCTCACTAAGCTCGTTGAGGTTGAGTATCT	780
Qy	257	LeuArgArgLeuAaSp	261
Db	781	CTACGCTCGTCTCGAT	795

```

RESULT 8
US-10-267-763-3
; Sequence 3, Application US/10267763
; Publication No. US20030077687A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Arabidopsis
US-10-267-763-3

```

Db	364	CTACGGCACAATATCCATCGCTTTGGACCCCTAAGGCTCTATCTGTTAAACATTATACAA	423
Qy	119	SerIleTrrpLysMetTyrPhelysGluIeuProIeuLeuAlaThrSerGlyAspAsp	138
Db	424	CAAAATCTGGGATATTTACTTTTAAAGAAATGCTTCCTTTGTTGTCAAACCTGCGCATGAT	483
Qy	139	GlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGlnSerIleSerArgArg	158
Db	484	GGCAACTATCCATCACTGCTGCTAGTAGTCTCGCCTGTTTACAGCTCTTTCGAAGAAG	543
Qy	159	IleHisTyrGlyLysPheValaGluValLysPheArgAspAlaProGlnAspTyrGlu	178
Db	544	ATTCACTACGGTAAATTTGTAGCTGAGGTCAAAATTCAGAGATGCTCCACAAGATTACGAG	603
Qy	179	ProLeuIleArgAlaLysAspLysGluClyLeuMetLysIeuIeuThrPheThrSerVal	198
Db	604	CTCGCAATTCGGCGCTCAGATGAGAGCGCTTGTATGAAGCTGTGACGTTTGAAGAAAGTA	663
Qy	199	GluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsn	218
Db	664	GAAGAAATCGTTTAAGACAGAGATCCAGAGAAGAACACAAACGTTTGGACAAGAGTAAAA	723
Qy	219	LeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe-----AspProSerValAla	236
Db	724	TTCAACTCTGGCTATGGCGATGAGAGTAAGAAGAAGTATAAAGTGGATCCATTGCTTGCC	783
Qy	237	SerSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeu	256
Db	784	TCTCGCATCTACGGGAATGCGTTATCCCTCTCACTAAGCTCTTGAGGTTCAGTATCTT	843
Qy	257	LeuArgArgLeuAsp	261
Db	844	CTACGCTCGTCTCGAT	858

```

RESULT 9
US-10-267-763-9/c
; Sequence 9, Application US/10267763
; Publication No. US20030077687A1
; GENERAL INFORMATION:
; APPLICANT: Grolach, Jörn
; APPLICANT: Boyss, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASES
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-267-763-9

Alignment Scores:
Pred. No.: 3, 07e-96
Score: 818.00
Percent Similarity: 72.83%
Best Local Similarity: 61.13%
Query Match: 61.09%
DB: 15
Length: 1006
Matches: 162
Conservative: 31
Mismatch: 68
Indels: 4
Gaps: 2

```

```
US-10-624-061-16 (1-261) x US-10-267-763-9 (1-1006)
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (432)
; NAME/KEY: unsure
; LOCATION: (514)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (516)
US-09-454-279-5
Alignment Scores:
Pred. No.: 4,31e-90
Score: 767.50
Percent Similarity: 95.29%
Best Local Similarity: 95.29%
Query Match: 57.32%
DB: 9
Gaps: 0
US-10-624-061-16 (1-261) x US-09-454-279-5 (1-525)
QY 1 MetAlaLysAlaAlaGluInSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20
Db 16 ATGCCAAAGCAGCAGCAACAAGTCTTGATCTGGGATGTGTACACGCTAGCTTCTGTG 75
QY 21 ArgGluAspLeuValArgGluInSerThrLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 76 AGAGAGGATTTGGTTAGGCAAGAGGATACCATCATCTTTATGGTCTCATTTGAGAGAGCCAA 135
QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnGlnGlnGlnGlnGln 60
Db 136 TTCCTTAGCAATTTCTCACCACTATGTAAGAAAGTATGCTCAATCCAGGGTTTGTGGC 195
QY 61 SerLeuValGluPheValValValValValValValValValValValValValVal 80
Db 196 TCATTGGTGGAAATTTGTTTAAAGATACAGAGGCCATTCAAGCTAAGGCTTGAAGATAC 255
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100
Db 256 AAAAACCCCTGAAGAAAGACGCCCTTTCTCCAGAAATTTACCACCATCAATTTGCCATCT 315
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120
Db 316 TACTCTTCAACACAGTTTTCGATCCCTGCTGCTTCAATTAACATAAACAAGTCATCT 375
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsn 140
Db 376 GGGAAATGTATTTCAAGAGAGTTACTTCCATTGCTTGTCTACTTCGGGTGATGATGGNAAC 435
QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgGileHi 160
Db 436 TATGCGCAAACTGCGCTATATGACCT-TCAATTATTGGAGTC-ATCTCTAGAGGAT-CA 492
QY 160 sTyrGlyLysPheValAlaGluValLys 169
Db 493 CTATGCAAAAGTTTGTAGCTGANGNGAAA 520
RESULT 11
US-10-624-061-5
; Sequence 5, Application US/10624061
; Publication No. US20040019929A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omelayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/10/624,061
US-09-454-279-5
; Sequence 5, Application US/09454279
; Publication No. US20020184658A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omelayo O.
; APPLICANT: Lee, Jian-Ming
; TITLE OF INVENTION: AROMATIC AMINO ACID BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB1299 US NA
; CURRENT APPLICATION NUMBER: US/09/454,279
; CURRENT FILING DATE: 1999-12-03
; EARLIER APPLICATION NUMBER: 60/110,845
; EARLIER FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 22
```

; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: US/09/454,279  
; PRIOR FILING DATE: 1999-12-03  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/110,845  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-04  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 5  
; LENGTH: 525  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (432)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (514)  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (516)  
US-10-624-061-5

Alignment Scores:  
Pred. No.: 4,31e-90 Length: 525  
Score: 767.50 Matches: 162  
Percent Similarity: 95.29% Conservative: 0  
Best Local Similarity: 95.23% Mismatches: 7  
Query Match: 57.32% Indels: 4  
DB: 16 Gaps: 0

US-10-624-061-16 (1-261) x US-10-624-061-5 (1-525)

```
QY 1 MetAlaLysAlaGluGlnSerProAspSerGlyAsnValThrLeuAlaSerVal 20
DB 16 ATGGCCAAAGCAGCAGACAAAGTCTGATTCGGGAATGTACACGCTAGCTTCGTG 75
QY 21 ArgGluAspLeuValArgGlnGluAspThrIleLeuValGlyLeuIleGluArgAlaLys 40
DB 76 AGAGAGGATTGGTTAGCAAGAGGATACCATCATTTATGCTCTCATTTAGAGAGCCAG 135
QY 41 PheProSerAsnSerHisThrTyraSpGluLysTyraAlaGlnIleGlnGlyPheCysGly 60
DB 136 TTCCTAGCAATTCACACCATATGATGAAAGATATGCTCAAAATCCAGGGTTTTTGTGGC 195
QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80
DB 196 TCATTGGTGAATTTGTTTGAAGATACAGAGCCATTCAGTAAAGCTGGAGATAC 255
QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100
DB 256 AAAAACCTGAAAGAAACGCTCTCTCCAGAAAATTTACCAACCAATTCATTCCTCATCT 315
QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnLysSerIle 120
DB 316 TACTCTTCAACAGTTTTCATCTCTGGTCTGCTTCATTAATACATAACAGTCACT 375
QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140
DB 376 GGGAAAATGTATTCAAAGATTACTTCCATTGCTTCTACTTCGGGTGATGAGNAAC 435
QY 141 TyrAlaGln-ThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHi 160
DB 436 TATGGCAAACTCAGCTAATGACCT-TCATTATTGAGTTC-ATCTCTAGAAGGAT-CA 492
QY 160 sTyGlyLysPheValAlaGluValLys 169
DB 493 CTATGGAAGTTTGTAGCTGANGNGAAA 520
```

## RESULT 12

US-09-938-842A-737

; Sequence 737, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAVE, AND METHODS OF USE  
; FILE REFERENCE: SCRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 737  
; LENGTH: 1005  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-737

Alignment Scores:  
Pred. No.: 1.37e-80 Length: 1005  
Score: 698.00 Matches: 131  
Percent Similarity: 68.08% Conservative: 46  
Best Local Similarity: 50.38% Mismatches: 75  
Query Match: 52.13% Indels: 8  
DB: 9 Gaps: 1

US-10-624-061-16 (1-261) x US-09-938-842A-737 (1-1005)

```
QY 10 AppSerGlyAsnValThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
DB 223 GAGAGAGTGAGAGTTTGACTCTCTGAAGGTATTAGAACTCTTTGATCCGTCAGAGGAC 282
QY 30 ThrIleLeuValGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyraSp 49
DB 283 AGCATATATTTGGGCTATTGGAGAGCCAGTACTGTTCATCATCTGTACTATTATGAT 342
QY 50 GluLysTyraAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
DB 343 CCTACTGCTTTTGACATGATGCTTCAATGCTTCTTGTGGTTGAGTACATGTTAAAGGC 402
QY 70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
DB 403 ACTGAGAGCTTCACGCTAAGTTGTTAGGTCTCTGATGATGAACATCTCTTCTTCTTC 462
QY 90 ProGluAsnLeuProProSerIleValProSerTyraSerPheLysGlnPheLeuHisPro 109
DB 463 CCTGATGATCTACAGAGCTATGTTGCTCTCTCTTCAAGTACCCAAAGCTGTGTGCAATTT 522
QY 110 GlyAlaAlaSerIleAsnIleAsnLysSerIleThrLysMetTyrPheLysGluLeuLeu 129
DB 523 GCTGCTGATTCGATAAACATAAACAAGAGATATGGAACATGTACTTCAGAGACCTTGTT 582
QY 130 ProLeuLeuAlaThrSerGlyAspGlyAsnTyraAlaGlnThrAlaAlaAsnAspLeu 149
DB 583 CCAAGACTTGGAAGAAAGCGATGATGTTAACTACGGCTCAACAGCTGTCTGTGACGCT 642
QY 150 SerLeuLeuGlnSerIleSerArgIleHisTyGlyLysPheValAlaGluValLys 169
DB 643 APTCCCTTCAGTGTCTCTCAAGAGAAATCCATACCGTAAATTTGTTGCAAGAGCTAAA 702
QY 170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
DB 703 TTTCAGGCTCACCGAAGCATACGAGTCCGCCATCAAGCAAGATAGAGTGCACCTG 762
QY 190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
DB 763 ATGGATATGCTGACATTCCTGAGATGCGAATCGGATAAAGAGAGAGATTGAGTGA 822
QY 210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAsnAsnGluAsnArg 229
```

```

Db      823 ACCGAAACATACCGGCAAGAGTGAAGTGGATGGAGGAGAAAGAAAGAGAA 882
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      883 GAAGGAATGAATCTCATGTTTACAAATACAGTCGATCTTAGTTGGTACATTATGGA 942
Qy      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261
Db      943 GATTGATCATGCCCTTTACAAAGAGGTTCAAGTGGAGTACTTCTGTCAGAGACTGGAC 1002

RESULT 13
US-09-938-842A-737
; Sequence 737, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 737
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-737

Alignment Scores:
Pred. No.: 1,37e-80 Length: 1005
Score: 698.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 52.13% Indels: 8
DB: 11 Gaps: 1

US-10-624-061-16 (1-261) x US-09-938-842A-737 (1-1005)
Qy      10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db      223 GATGAGAGTGAGAGTTTGACTCTTGAAGGTATTAGAACTCTTTGATCGGTCAAGAGGAC 282
Qy      30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
Db      283 AGCATATTATTTGGGTATTGGAGAGCCCAAGTACTGTTTACAACTGCTGATCTATTGAT 342
Qy      50 GluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValLysAsn 69
Db      343 CCTACTCTTTTCACATGGATGTTTCAATGGTCTCTTGGTTCAGTACATGGTTAAAGGC 402
Qy      70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db      403 ACTGAGAAGCTTCACGCTAAGTTGGTAGTTTAAAGAGTCTCTGATGAACATCTCTTCTTC 462
Qy      90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db      463 CCTGATGATCTTACAGAGCCTATTGTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 522
Qy      110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
Db      523 GCTGCTGATTCGATAACATAACAAAGAGATATGGAACATGATGATCTTCAGAGACCTTGT 582
Qy      130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149

```

```

Db      583 CCAGAGCTTTGTGAAGAAAGCGGATGATGGTAATCACTACGGCTCAACAGCTGTCTGTGAGCT 642
Qy      150 SerLeuLeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db      643 ATTCGCCCTTCAGTGTCTCTCAAGAGAAATCCATTACGTAATTTGTTCGAGAAGCTAAA 702
Qy      170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      703 TTTCAAGCCCTCACCCGGAAGCATACGAGTCGCGCATCAACAGCAAGATAGGATGCACTG 762
Qy      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      763 ATGGATATGTCGATCTCCGACTGTGGAAGATGCGATAAAGAGAGAGTTGAGATGAAA 822
Qy      210 AlaValValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAspAsnGluAsnArg 229
Db      823 ACCCGAATACATACGGGCAAGAGTGAAGTTGGGATGGAGAGAAAGAAAGAAAGAA 882
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      883 GAAGGAATGAATCTCATGTTTACAAATACAGTCCGATCTTAGTTGCTGACTTATATGGA 942
Qy      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeuAsp 261
Db      943 GATTGATCATGCCCTTTACAAAGAGGTTCAAGTGGAGTACTTGTCTCAGAGACTGGAC 1002

RESULT 14
US-10-267-763-2
; Sequence 2, Application US/10267763
; Publication No. US20030077687A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1207
; TYPE: DNA
; ORGANISM: Arabidopsis
US-10-267-763-2

Alignment Scores:
Pred. No.: 3,32e-80 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 68.08% Conservative: 46
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.99% Indels: 8
DB: 15 Gaps: 1

US-10-624-061-16 (1-261) x US-10-267-763-2 (1-1207)
Qy      10 AspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArgGlnGluAsp 29
Db      235 GATGAGAGTGAGAGTTTGACTCTTGTGAAGTATTGAAACTCTTGTATCCGTCAGAGGAC 294
Qy      30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49

```

```

Db      295 AGCAATTATATTGGGCTATTGGAGAGAGCCAAAGTACTGTTCACATGCTGTAFACTTTATGAT 354
Qy      50 GluLysTyrAlaGlnleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
Db      355 CCTACTGCTTTTCACATGATGGTTTCATGATGTTCTTTGGTTGAGTACATGCTTAAAGGC 414
Qy      70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db      415 ACTGAGAGCTTTCACGCTAAGGTTGGTGGTATTTAAGAGTCTCTGATCAACATCCCTTTCTTC 474
Qy      90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db      475 CCTGATGATCTACAGAGCTATGTTGCTCTCTTCAGTACCCAAAGGTGTGCAATTT 534
Qy      110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGlnLeuLeu 129
Db      535 GCTGCTGATTCGATAAACAATAACAAGAGATATGGAACATGTACTTCAGAGACCTTGTT 594
Qy      130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
Db      595 CCAAGACTGTGAAGAGCGATGATGTAAGTAACTACGCTCAACAGCTGTCTGTGACGCT 654
Qy      150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db      655 ATCTGCTTTCAGTCTCTCAAGAGAAATCCATTACGGTAAATTTGTCAGAGAGCTAAA 714
Qy      170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      715 TTTCAGGCTCACCGAGCATACGAGTCGCGCATCAAGACACAGATAAGATGCGACTG 774
Qy      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      775 ATGGATATGCTGACATTCGACTCCGACTCGAAGATGCGATAAAGAGAGAGATTGAGATGAA 834
Qy      210 AlaValValPheGlnGluValAsnSerAspAsnSerAspAsnAspAsnGluAsnArg 229
Db      835 ACCCGAACATACGGGCAAGAGTGAAGTTGGATGGAGAGAAAGAGAGAGAGAA 894
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      895 GAAGGGAATGATCTCATGTTTACAAATCATGTCGATCTTAGTTGGTACTTATATGGA 954
Qy      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
Db      955 GATTGGATCATGCTTTTAAACAAAGAGGTTCAAGTGGAGTACTTTGCTCAGAAGACTGGAC 1014

RESULT 15
; US-10-267-763-8/c
; Sequence 8, Application US/10267763
; Publication No. US20030077687A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: Boyes, Douglas
; APPLICANT: Davis, Keith
; APPLICANT: Hamilton, Carol
; APPLICANT: Hoffman, Neil
; APPLICANT: Kloti, Andreas
; APPLICANT: Woessner, Jeffrey
; APPLICANT: Zayed, Adel
; APPLICANT: Ascenzi, Robert
; TITLE OF INVENTION: MODULATION OF CHORISMATE SYNTHASE AND CHORISMATE MUTASE
; TITLE OF INVENTION: EXPRESSION OR ACTIVITY IN PLANTS
; FILE REFERENCE: 2035DIV1
; CURRENT APPLICATION NUMBER: US/10/267,763
; PRIORITY FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 09/610,040
; PRIOR FILING DATE: 2000-07-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1207
; TYPE: DNA

```

```

; ORGANISM: Arabidopsis thaliana
US-10-267-763-8

```

```

Alignment Scores:
Pred. No.: 3 32e-80 Length: 1207
Score: 696.00 Matches: 131
Percent Similarity: 68.08% Conservative: 75
Best Local Similarity: 50.38% Mismatches: 75
Query Match: 51.98% Indels: 8
DB: 15 Gaps: 1

```

```

US-10-624-061-16 (1-261) x US-10-267-763-8 (1-1207)

```

```

Qy      10 AspSerGlyAsnValTyrThrLeuAlaSerValArgLysLeuValArgGlnGluAsp 29
Db      973 GATGAGAGTGAGAGTTTGAATCTCTTGAAGGTATTAGAAACTCTTTGATCCGTCAGAGGAC 914
Qy      30 ThrIleIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAsp 49
Db      913 AGCATTTATATTGGGCTATTGAGAGAGCCAAAGTACTGTTACATGCTCATATTATGAT 854
Qy      50 GluLysTyrAlaGlnleGlnGlyPheCysGlySerLeuValGluPheValValLysAsn 69
Db      853 CCTACTGCTTTTGACATGATGTTTCAATGGTCTTTGGTTGAGTACATGTTTAAAGGC 794
Qy      70 ThrGluAlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePhe 89
Db      793 ACTGAGAGCTTTCACGCTAAGGTTGGTGGTAAAGTCTCTGATGAACATCCCTTTCTTC 734
Qy      90 ProGluAsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisPro 109
Db      733 CCTGATGATCTACAGAGCTATGTTGCTCTCTTCAGTACCCAAAGGTGTGCAATTT 674
Qy      110 GlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeu 129
Db      673 GTGCTGATTCGATAAACAATAACAAGAGATATGGAACATGTACTTTCAGAGACCTTGT 614
Qy      130 ProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeu 149
Db      613 CCAAGACTGTGAAGAGCGATGATGTTAACTACGGCTCAACAGCTGCTGTGACGCT 554
Qy      150 SerLeuLeuGlnSerIleSerArgIleHisTyrGlyLysPheValAlaGluValLys 169
Db      553 ATCTGCTTTCAGTCTCTCAAGAGAAATCCATTACGGTAAATTTGTTGCAAGAGCTAAA 494
Qy      170 PheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeu 189
Db      493 TTTCAGGCTCACCGAGCATACGAGTCCGCCATCAAGACACAGATPAAGATCGACTG 434
Qy      190 MetLysLeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLys 209
Db      433 ATGGATATGCTGACATTCGACTTCGAGAGATGCGATAAAGAGAGAGATTGAGATGAAA 374
Qy      210 AlaValValPheGlnGluValAsnSerAspAsnSerAspAsnAspAsnGluAsnArg 229
Db      373 ACCCGAACATACGGGCAAGAGTGAAGTTGGATGGAGAGAAAGAGAGAGAGAGAA 314
Qy      230 -----LysPheAspProSerValAlaSerSerLeuTyrLys 241
Db      313 GAAGGGAATGATCTCATGTTTACAAATCATGTCGATCTTAGTTGGTACTTATATGGA 254
Qy      242 AsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgArgLeuAsp 261
Db      253 GATTGGATCATGCTTTTAAACAAAGAGGTTCAAGTGGAGTACTTCTCAGAAGACTGGAC 194

```

Search completed: September 25, 2004, 06:20:33

Job time : 520 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 25, 2004, 02:58:27 : Search time 2777 Seconds  
(without alignments)  
2806.636 Million cell updates/sec

Title: US-10-624-061-16  
Perfect score: 1339  
Sequence: 1 MAKAAEQPSDGNVTLASV.....NWVPLTKVQVYLLRLD 261

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_p2n.model -DRV=xlh  
-Q=/cgm2\_1/USPTO.spool/US10624061/runat\_24092004.092407.16924/app\_query.fasta.1.455  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORW=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10624061@cgm2\_1.3437@runat\_24092004.092407.16924 -NCPU=6 -ICPU=3  
-NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:  
2: em\_esthum:  
3: em\_estin:  
4: em\_estov:  
5: em\_estpl:  
6: em\_estro:  
7: em\_esttc:  
8: gb\_est1:  
9: gb\_est2:  
10: gb\_est3:  
11: gb\_est4:  
12: gb\_est5:  
13: gb\_est6:  
14: gb\_est7:  
15: em\_estfun:  
16: em\_estom:  
17: em\_gss\_hum:  
18: em\_gss\_inv:  
19: em\_gss\_pln:  
20: em\_gss\_vrt:  
21: em\_gss\_fun:  
22: em\_gss\_mam:  
23: em\_gss\_mus:  
24: em\_gss\_pro:  
25: em\_gss\_rod:  
26: em\_gss\_phg:  
27: em\_gss\_vrl:  
28: gb\_gss1:

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	917	68.5	563	10	AW666427 sk36d12.y
2	895	66.8	648	10	AW774625 EST333776
3	876	65.4	512	12	BG363365 sac13f08.
4	831	62.1	654	10	AW775842 EST334907
5	831	62.1	696	12	B1273174 NF091812F
6	794	59.3	605	10	BF649665 NF081H11E
7	755	56.4	620	10	BF009709 sp83b10.y
8	723.5	54.0	605	12	B1420527 LJNEST58b
9	721	53.8	477	9	A1973566 sc88e02.y
10	721	53.8	560	10	BE821647 GM700015A
11	717	53.5	529	14	CB827129 LJNEST70f
12	702	52.4	422	12	BG239446 sab70d03.
13	679.5	50.7	810	14	CD427426 SA1_30_H0
14	672	50.2	560	12	EM093460 saj09c02.
15	671	50.1	1196	11	AY107907 Zsa may8
16	670.5	50.1	1222	11	AY103806 Zsa may8
17	670	50.0	748	14	CD824727 BN25.058U
18	667	49.8	726	14	CA800612 sau19a05.
19	661.5	49.4	710	14	CF475712 RTW2.11
20	649.5	48.5	792	12	BG309831 HVSMEC001
21	622.5	46.5	660	13	BX253798 BX253798
22	618.5	46.2	821	14	CK196770 FGAS00523
23	609.5	45.5	867	14	CB619882 OS11a04L
24	608.5	45.4	738	13	BQ989273 QGF17D13.
25	604.5	45.1	524	14	CB828627 LJNEST88g
26	590.5	44.1	531	14	CB828445 LJNEST86f
27	590.5	44.1	557	14	CB829043 LJNEST93e
28	586.5	43.8	516	14	CB827444 LJNEST74h
29	583.5	43.6	680	9	AJ612241 AJ612241
30	580	43.3	885	13	BU693059 SYD0085 P
31	579.5	43.3	761	14	CF480460 POL1_65_H
32	579	43.2	639	10	BE660542 2-B12 Gma
33	578.5	43.2	501	14	CB828062 LJNEST82C
34	573.5	42.8	736	14	CF443707 EST680052
35	573	42.8	698	10	BE577489 L48-2335T
36	572.5	42.8	549	14	CB828614 LJNEST88f
37	571.5	42.7	529	14	CB826768 LJNEST66f
38	571	42.6	680	13	CA080410 SCVPAM105
39	567.5	42.4	541	14	CB829238 LJNEST95g
40	564	42.1	683	14	CF481222 POL1_70_E
41	563.5	42.1	725	14	CF486896 POL1_40_B
42	561	41.9	696	14	CD919148 G608_112E
43	559	41.7	595	10	BE445201 WHE1133G
44	552.5	41.3	830	14	CB622596 OS11a091
45	550.5	41.1	711	10	BF430733 OG04A12T3

ALIGNMENTS

RESULT 1  
AW666427  
LOCUS  
DEFINITION  
AW666427 563 bp mRNA linear EST 03-DEC-2001  
sk36d12.y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-cl028-4272 5', similar to SM:CHMU\_ARATH P42738 CHORISMATE MUTASE  
PRECUSOR ; mRNA sequence.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
AW666427 GI:7478842  
Glycine max (soybean)  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;



rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1 (bases 1 to 563)

**REFERENCE**  
**AUTHORS**  
 Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Aucuba, T., Martin, J., Beck, C., Wylie, J., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, J., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
 Public Soybean EST Project  
 Unpublished (1999)

**TITLE**  
**JOURNAL**  
**COMMENT**  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com  
 High quality sequence stop: 418.

# FEATURES

source

1. 563  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1028-4272"  
 /tissue\_type="roots of 'Supernod' plants"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-c1028"  
 /note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; The mRNA was isolated from roots of Glycine max 'Supernod' plants generously donated by Dr. Gary Stacey. The seedlings were inoculated with Bradyrhizobium japonicus, strain USDA110 prior to harvest. Stratagene's cDNA synthesis kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A,C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAGAGAGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second-strand synthesis, the cDNA ends were filled in with cloned Pfu DNA polymerase, ligated to EcoRI adapters and subsequently phosphorylated. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.16e-103 Length: 563  
 Score: 917.00 Matches: 176  
 Percent Similarity: 97.85% Conservative: 6  
 Best Local Similarity: 94.62% Mismatch: 4  
 Query Match: 68.48% Indels: 0  
 DB: 10 Gaps: 0

US-10-624-061-16 (1-261) x AW666427 (1-563)

Qy 25 ValArgGlnGluAspThrIleTyrGlyLeuIleGluArgAlaLysPheProSerAsn 44

Db 6 GTTAGGCAAGAGGATACCATCATTTATGGTCTCATTTAGAGAGCCAAAGTTCCCTAGCAAT 65

Qy 45 SerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGlySerLeuValGlu 64

Db 66 CCTCACACCTATCATGAAGAGTATGCTCAATCCAGGGTTTTGTGGCTCATTTGGTAGAA 125

Qy 65 PheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyValTyrLysAsnProGlu 84

Db 126 TTGTGTGTAGATACAGAGCCATTCAGCTGAGCTGGAGATACAAAACCTGAA 185

Qy 85 GluAsnAlaPhePheProGluAsnLeuProProSerIleValProSerTyrSerPheLys 104

Db 186 GAAAGCGCTTCTCCAGAAAAATTTACCACCATCAATTTGTGCTATCTTACTCTCTTCAA 245

Qy 105 GlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIleTyrLysMetTyr 124

Db 246 CAGTTTTTGATCCTCGAGCTGCTTCAATTACCAATTAACAAGTCCATCTGGAAATGTAT 305

Qy 125 PheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThr 144

Db 306 TTCCAAGAGTTACTTCCATTGTTGCTACTTCGGGGGATGATGGAAACTATGCACAACT 365

Qy 145 AlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgArgIleHisTyrClyLysPhe 164

Db 366 GCAGCTAAATGATCTTTCAATTGTCAGGCGCATCTCTAGAAGGANTTCACTATGAAAGTTT 425

Qy 165 ValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeuIleArgAlaLys 184

Db 426 GTAGCTGAGGTGAATTCANGATGCTCTCAAGACTACGAGCCTTTTATTTCGAGCTAAG 485

Qy 185 AspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGluThrValArgLys 204

Db 486 GATAAAGAGAGGATTGATGAATTTGGTGATTCATTACAGCGTTGAAGAGACAGTGAGGAA 545

Qy 205 ArgValGluLysLysAla 210

Db 546 GAGGTGAAGAAGAGGCC 563

## RESULT 2

AW774625 648 bp mRNA linear EST 07-SEP-2000  
 EST333776 KV3 Medicago truncatula cDNA clone pkV3-23M13, mRNA  
 sequence.  
 ACCESSION AW774625  
 VERSION AW774625.1 GI:7718542  
 KEYWORDS EST.  
 SOURCE Medicago truncatula (barrel medic)  
 ORGANISM Medicago truncatula  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 1 (bases 1 to 648)  
 VandenBosch, K., Hurt, J., Moore, J., Beremand, P., Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B., Hansen, T.S., Holt, I.E. and Fraser, C.M.  
 ESTs from roots of Medicago truncatula after Rhizobium inoculation  
 Unpublished (1999)  
 Contact: VandenBosch K  
 Department of Plant Biology  
 University of Minnesota  
 220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
 Tel: 612 624 2755  
 Fax: 612 625 1738  
 Email: kvandenbosch@umn.edu  
 Texas A&M EST name: T258044e  
 TIGR sequence name: MTEBC79TK  
 More information is available at:  
 http://chrysie.tamu.edu/medicago  
 Seq primer: SKmod (CTA GAA CTA gfg gAT CC).  
 Location/Qualifiers  
 1. 568

## FEATURES

source



```

/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKV3-23M13"
/tissue_type="Seedling roots"
/dev_stage="3 days post-inoculation with Sinorhizobium
melliotti"
/lab_host="E. coli strain XLOLR"
/clone_lib="KV3"
/notes="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unisap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant Lambda-Zap phage using Ex-assist
helper phage and propagated in XLOR cells."

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 8,02e-101 Length: 648
Score: 895.00 Matches: 167
Percent Similarity: 87.50% Conservative: 22
Best Local Similarity: 77.31% Mismatches: 27
Query Match: 66.84% Indels: 0
DB: 10 Gaps: 0

```

US-10-624-061-16 (1-261) x AW774625 (1-648)

```

Qy 34 GlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyrAspGluLysTyrAla 53
Db 1 GGTCTATTGAGAGATCAAAAGTTCCCTTAATTCACAACTATGATCAGAAATTATCT 60
Qy 54 GlnIleClnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAlaIle 73
Db 61 CAAATCCCTGGTCTTTGTGGCTCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
Qy 74 GlnAlaLysAlaGlyArgTyrLysAsnProGluGluAlaPhePheProGluAsnLeu 93
Db 121 CAGCTAAGCTGGAAGATACAAAACCTGAGAGATCCCTTCCTCCAGAAATTTA 180
Qy 94 ProProSerIleValProSerTyrSerPheGlyGlnPheLeuHisProGlyAlaAlaSer 113
Db 181 TCATTGTCACTTGTTCCTTACCCCTTCACAAAGTTTGTGATCCTCTGGAGCCGCTCG 240
Qy 114 IleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLysLeuAla 133
Db 241 ATTAACATAACAAAGTCTTATGAAATGATCTTTCATGACCTGCTTCATTTGTCGT 300
Qy 134 ThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGln 153
Db 301 GCCTCCGGTGATGATGCAACTATGCAAACTGCAGTAGCAGCCTCTCATTTGTCAG 360
Qy 154 SerIleSerArgAlaGlnHisTyrGlyLysPheValAlaGluValLysPheArgAspAla 173
Db 361 GCCATTTCTAAAGAGTTCATTATGGAAGTTGTAGCTGAAGTGAATTCAGGAATCT 420
Qy 174 ProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeu 193
Db 421 CCTCAAGACTATGACCTTAATTCGCTCAGGACAAAGAGGTTGATGAATTTG 480
Qy 194 ThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaValAlaPhe 213
Db 481 ACATTTAAGAGTGTGAAGAGATGTTAAAGAAAGGGTTGAAAGAGGAGGAAATGTTT 540
Qy 214 GlyGlnGluValAsnLeuAsnSerAspAsnAspAsnGluAsnArgLysPheAspPro 233
Db 541 GGGCAGGATATPACCTTACAGCAGTGTGACGTAAAGGTAAACAAAGTTTGTATCCA 600
Qy 234 SerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThrLys 249
Db 601 TCAGTGGCTTCTAAGTGTATGAGAAATGGGTATACCTCTTACTAAG 648

```

RESULT 3  
BG363365  
LOCUS

DEFINITION

BG363365 512 bp mRNA linear EST 28-NOV-2001  
sac19f08.y1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE  
ID: Gm-cl051-2943 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE  
; mRNA sequence.

ACCESSION

VERSION BG363365.1 GI:13252454

SOURCE

ORGANISM Glycine max (soybean)

KEYWORDS

Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

REFERENCE

AUTHORS 1 (bases 1 to 512)

Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Coryell,V.,  
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
McCann,R., Waterston,R. and Wilson,R.

TITLE

JOURNAL

COMMENT

Public Soybean EST Project

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130

South Memorial Parkway Huntsville, AL 35801 For further information

call: (800)-533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 397.

FEATURES

source

1..512

/organism="Glycine max"

/mol\_type="mRNA"

/db\_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-2943"

/tissue\_type="floral meristematic mRNA"

/lab\_host="DH10B"

/clone\_lib="Gm-cl051"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; The cDNA library was constructed from floral

meristematic mRNA provided by Dr. Halina Knap of Clemson

University. Complementary DNA was synthesized from mRNA

using a primer consisting of a poly(dT) sequence with a

XhoI restriction site. EcoRI adapters were ligated to the

blunt-ended cDNA fragments followed by XhoI digestion. The

cDNA fragments were directionally cloned into the

EcoRI-XhoI restriction site of the pBluescript vector. The

ligated cDNA fragments were transformed into DH10B host

cells (GibcoBRL). This library was constructed in the

laboratory of Dr. Randy Shoemaker."

## ORIGIN

Alignment Scores:

Pred. No.: 1.32e-98 Length: 512

Score: 876.00 Matches: 169

Percent Similarity: 99.41% Conservative: 0

Best Local Similarity: 99.41% Mismatches: 1

Query Match: 65.42% Indels: 0

DB: 12 Gaps: 0

US-10-624-061-16 (1-261) x BG363365 (1-512)

Qy 7 GlnSerProAspSerGlyAsnValTyrThrLeuAlaSerValArgGluAspLeuValArg 26

Db 3 CAAAGTCTTGATTCGGGAATGTGACACGTAGCTTCTTGAGAGAGATTGTTAGG 62

Qy 27 GlnGluAspThrIleTyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHis 46

Db 63 CAAGAGGATACCATCATTTATGCTCTCTATGAGAGGCAAGTTCCTAGCAATTTCTCAC 122  
 QY ThrTyAspGluLysTyAlaGlnGlnGlyPheCysGlySerLeuValGluPheVal 66  
 Db 123 ACCTATGATCAAAAGTATGCTCAATCAACAGGGTTTTGTGGCTCATTTGGTGAATTTGTT 182  
 QY ValLysAsnThrGluAlaGlnGlnLysAlaGlyArgTyLysAsnProGluGluAsn 86  
 Db 183 GTTAAGAATACAGAGGCCATTCAGCTAGGCTGGAAGATACAAAAACCTCAAGAAAC 242  
 QY AlaPhePheProGluAsnLeuProSerIleValProSerTySerPheLysGlnPhe 106  
 Db 243 GCTTCTCCAGAAATTTACACCAATGTCATGTCATCTTCTTCAACAGTTT 302  
 QY LeuHisProGlyAlaAlaSerIleAsnLysSerIleTrpLysMetTyPheLys 126  
 Db 303 TTGCATCTCTGGTCTCTCAATTAACATAAACAAGTCCATCTGGAAATGTATTTCAA 362  
 QY GluLeuLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyAlaGlnThrAla 146  
 Db 363 GAGTTACTTCTCCATTTCTCTACTTCTGGGTGATGCGCAACTATGCCAAACCTGCAGCT 422  
 QY AsnAspLeuSerLeuLeuGlnSerIleSerArgGlnGlnHisTyLysGlyPheValAla 166  
 Db 423 AATGACCTTTCTATTTGTCAGTCCATCTCTAGAGGATTCATGGAAGTTTGTAGCT 482  
 QY GluValLysPheArgAspAlaProGlnAsp 176  
 Db 483 GAGTGAATTCAGGAGTCTCTCTCAAGAC 512

RESULT 4  
 AW775842  
 LOCUS  
 DEFINITION EST334907 DSIL Medicago truncatula cDNA clone pDSIL-3E8, mRNA  
 ACCESSION AW775842  
 VERSION AW775842.1 GI:7765655  
 KEYWORDS  
 SOURCE Medicago truncatula (barrel medic)

ORGANISM  
 Eukaryota; Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;  
 Medicago.  
 1 (bases 1 to 654)  
 Pedorova, M., Pierson, B.L., Samac, D.A., Vance, C.P., Gantt, G.S.,  
 Peng, H., Ellis, L., Town, C.D., Bowman, C.L., Craven, M.B.,  
 Hansen, F.S., Holt, I.B. and Fraser, C.M.  
 ESTs from leaves of Medicago truncatula after inoculation with  
 Colletotrichum trifolii  
 Unpublished (2000)  
 Contact: Deborah A. Samac  
 Department of Plant Pathology  
 University of Minnesota  
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA  
 Tel: 612 625 1243  
 Fax: 612 649 5058  
 Email: debbys@puccini.crl.umn.edu  
 Minnesota sequence name: M259175e  
 TIGR sequence name: MTFa728TK  
 More information is available at:  
 http://chrysis.tamu.edu/medicago  
 Seq primer: SKmod (CTA GAA CTA Gtg GAT CC).  
 Location/Qualifiers  
 1..654  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultiVar="genotype Al7"  
 /db\_xref="taxon:3880"  
 /clone="pDSIL-3E8"  
 /tissue type="leaves infected with Colletotrichum trifolii"

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

FEATURES  
 source  
 1..654  
 /organism="Medicago truncatula"  
 /mol\_type="mRNA"  
 /cultiVar="genotype Al7"  
 /db\_xref="taxon:3880"  
 /clone="pDSIL-3E8"  
 /tissue type="leaves infected with Colletotrichum trifolii"

RESULT 5  
 BI273174  
 LOCUS  
 DEFINITION NF091B12FL1F1096 Developing flower Medicago truncatula cDNA clone  
 ACCESSION NF091B12FL 5', mRNA sequence.  
 BI273174  
 VERSION BI273174.1 GI:14883109

/dev\_stage="cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii"  
 /lab\_host="E. coli strain XLOLR"  
 /clone\_lib="DSIL"  
 /notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA from cotyledons and primary leaves harvested 5 and 8 days after inoculation with Colletotrichum trifolii. The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in XLOLR cells. Note: EST may be of fungal origin."

## ORIGIN

Alignment Scores:  
 Pred. No.: 7,92e-93 Length: 654  
 Score: 831.00 Matches: 157  
 Percent Similarity: 88.83% Conservative: 18  
 Best Local Similarity: 79.70% Mismatches: 22  
 Query Match: 62.06% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-624-061-16 (1-261) x AW775842 (1-654)  
 QY 13 AsnValTyThrLeuAlaSerValArgGluAspLeuValArgGlnAspThrIle 32  
 Db 64 AATGAATATATCTCTTGATCAGTGAGGAGACTTTGGTTAGACAAGATCAATTTGTT 123  
 QY 33 TyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyAspGluLysTy 52  
 Db 124 TTTGGTCTTATTGAGAGATCAAGATTTCTTTTAATTTCTCAAACTTATGATCAGAATTAT 183  
 QY 53 AlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAla 72  
 Db 184 CTTCAATCCCTGGTTTTTGTGGCTCATTTGGTTGAGTTTGTTCACCAATCTCAGATT 243  
 QY 73 IleGlnAlaLysAlaGlyArgTyLysAsnProGluGluAsnAlaPhePheProGluAsn 92  
 Db 244 GTTCAAGCTAAGCTGGAAGATACAAAACCTCGAAGAGATCCCTCTTCCAGAAAT 303  
 QY 93 LeuProProSerIleValProSerTySerPheLysGlnPheLysHisProGlyAlaAla 112  
 Db 304 TTATCATTTGCTCATTTGTTCCATCTTACCCCTTACAAAAGTTTTTGCATCTCTGAGCCGCT 363  
 QY 113 SerIleAsnIleAsnLysSerIleTrpLysMetTyPheLysGluLeuLeuProLeuLeu 132  
 Db 364 TCGATTACATTAACAAAGTCTTATGGAATAAGTACTTTTCATGACTGCTTCCATTTGTC 423  
 QY 133 AlaThrSerGlyAspAspGlyAsnTyAlaGlnThrAlaAlaAsnAspLeuSerLeuLeu 152  
 Db 424 GTTGCTTCCGGTGATGATGGCAACTATGCACAACTGCAGCTAGCGACCTCTCATTTGTT 483  
 QY 153 GlnSerIleSerArgArgIleHisTyGlyLysPheValAlaGluValLysPheArgAsp 172  
 Db 484 CAGGCCATTTCTAAAGAGTTTCAATTTGGAAGTTTGTAGCTGAAGTGAATTCAGGAA 543  
 QY 173 AlaProGlnAspTyGluProLeuIleArgAlaLysAspLysGluGlyLeuMethLysLeu 192  
 Db 544 TCTCCTCAAGACTATGAGCCCTTTAATTCGGTCTAAGGACAAAGAGGGTTGATGAATTTG 603  
 QY 193 LeuThrPheThrSerValGluGlnThrValaGlyLysArgValGluLysLys 209  
 Db 604 TTGACATTTAAGAGTGTGGAAGAGATGTTAAAGAAAGGGTTGAAAGAAAG 654

```

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 696)
Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula flower library
Unpublished (2001)
Contact: May GD
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert length: 696 Std Error: 0.00
Plate: 091 row: B column: 12
Seq primer: TCACACAGGAACAGCTATGAC.

FEATURES
source
1..696
/organism="Medicago truncatula"
/mol_type="mRNA"
/db_xref="taxon:3880"
/clone="NF091B12FL"
/tissue_type="Developing flowers"
/dev_stage="developmentally pooled. Contains a mixture of
very young, developing, fully-opened flowers and
flowers in early transition into pods."
/clone_lib="Developing flower"
/notes="Vector: lambda Zap; cDNA was prepared from polyA+
enriched, pooled samples of equivalent amounts of total
RNA from very young, developing, fully-opened flowers and
flowers transitioning into pods. The cDNA was
directionally ligated into the Uni-Zap XR vector
(Stratagene) and packaged using the Gigapack III Gold
packaging extracts. Phagemids containing cDNA inserts were
in vivo excised from the recombinant Uni-Zap XR vector
using ExAssist helper phage and the E. coli strain
XL1-Blue MRF' (Stratagene). Excised plasmids were plated
using SOLR cells."
```

ORIGIN

Alignment Scores: 8.71e-93 Length: 696

Pred. No.: 831.00 Matches: 159

Score: 86.47% Conservative: 20

Percent Similarity: 76.81% Mismatches: 27

Best Local Similarity: 62.06% Indels: 1

Query Match: 12 Gaps: 0

DB:

US-10-624-061-16 (1-261) x BI273174 (1-696)

QY 52 TyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValVallysAsnThrGlu 71

Db 5 TATCTCAATCCCTGGCTTTTGGCTCATCTGGTTGAGTTTGTTCACCAATCTGAG 64

QY 72 AlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGlnAsnAlaPheProGlu 91

Db 65 ATTGTTCAAGTAAAGGTGGAAGATACAAAACCTGAAGAGAAATCCCTCTTTCAGAA 124

QY 92 AsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyVala 111

Db 125 AATTATCATGTGCTACTGTTCATCTTACCCCTTCAAGATTGTCATCTGGAGCC 184

QY 112 AlaSerIleAsnLysSerIleTrpLysVetTyrPheLysGluLeuLeuProLeu 131

Db 185 GCTTCGATTAACTAAACAGTCTTATGGAATGTACTTTTCATGACCTGCTTCATTG 244

QY 132 LeuAlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerIleu 151

Db 245 TTCGTGCTTCGGTGATGATGGCAACTATGCACAACTGCAGTAGCGACTCTCATTT 304

QY 152 LeuGlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArg 171

Db 305 TTGCAGGGCACTTTTAAAGAGTTCATTATGAAAAGTTTGTAGCTGAAGTGAATTCAGG 364

QY 172 AspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLys 191

Db 365 GAATCTCTCAAGACTATGAGCCTTTAATTCGGTCTAAGGACAAAGAGGGTTGATGAAA 424

QY 192 LeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaVal 211

Db 425 TTGTTGACATTTAAGAGTGTTCAGAGAGATGCTAAGAAAAGGGTTGAAAAGAGGCAACA 484

QY 212 ValPheGlyGlnGluValAsnLeuAsnSerAspAspAsnAspAsnGluAsnArgLysPhe 231

Db 485 TTGTTTGGGCAGGATATTAAGCCTTAACACAGCAGTGTATGACAGTAAAGGTTAAACAAGTTT 544

QY 232 AspProSerValAlaSerSerLeuTyrLysAsnTrpValIleProLeuThr-LysGluVa 251

Db 545 GATCCATCAGTGGCTTCTAAGTTGTATGANAATGGGTAACTACCTCTTACTAANGANGGT 604

QY 251 lclnValGluTyrLeuLeu 257

Db 605 CAAGGTTGAGTACTACTG 623

RESULT 6

BF649665 605 bp mRNA linear EST 20-DEC-2000

LOCUS

DEFINITION

clone NF081H11EC 5', mRNA sequence.

ACCESSION

BF649665

VERSION

BF649665.1 GI:11914795

KEYWORDS

Medicago truncatula (barrel medic)

SOURCE

Medicago truncatula

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE

1 (bases 1 to 605)

Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.

Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research

Unpublished (2000)

Contact: Dixon RA

Plant Biology Division

The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7302

Fax: 580 221 7380

Email: radixon@noble.org

Insert Length: 605 Std Error: 0.00

Plate: 081 row: H column: 11

Seq primer: TCACACAGGAACAGCTATGAC.

Location/Qualifiers

1..605

/organism="Medicago truncatula"

/mol\_type="mRNA"

/db\_xref="taxon:3880"

/tissue\_type="Cell cultures derived from root tissues"

/dev\_stage="Cell suspensions were subcultured every 14 days. Cells were induced six days after subculture"

/clone\_lib="Elicited cell culture"

/note="Vector: Lambda Zap; Cells were induced with yeast cell wall extracts equivalent to 50ug/ml glucose in the final concentration. Samples were taken at 0.5, 1, 12 and 24 hours after induction. Equal amounts of RNA from each time point were pooled and used for mRNA isolation."

## ORIGIN

Alignment Scores: 2,92e-88 Length: 605  
 Pred. No.: 794.00 Matches: 150  
 Score: 87.50% Conservative: 13  
 Percent Similarity: 78.12% Mismatches: 24  
 Best Local Similarity: 59.30% Indels: 0  
 Query Match: 10  
 DB: 10

US-10-624-061-16 (1-261) x BF649665 (1-605)

QY 13 AnValTyThrLeuAlaSerValArgGluAspLeuValArgGlnGluAspThrIleLe 32  
 Db 29 ATGAATATACCTTGTATACGTGAGGAGACTTTGGTTAGACAGAGATACAAATTGT 88  
 QY 33 TyrGlyLeuIleGluArgAlaLysPheProSerAsnSerHisThrTyAspGluLysTyr 52  
 Db 89 TTGGTCTTATTGAGAGATCAAGTTCTCTTTAATCTCAACCTTATGATCAGAAATTAT 148  
 QY 53 AlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAla 72  
 Db 149 CTTCAATACCCCGTTTGTGGCTCATTTGGTGTGATTTGTTTCCCAATACTGAGATT 208  
 QY 73 IleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsn 92  
 Db 209 GTTCAAGCTTAAGCTGGAGATACAAACCTTGAGAGATCCCTTCTTCCAGAAAT 268  
 QY 93 LeuProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAla 112  
 Db 269 TTATCATTTGCTACTTGTCTTCACTTACCCCTTCAAAAGTTTTCATCTCGAGCCGCT 328  
 QY 113 SerIleAsnIleAsnLysSerIleTyrPheLysGluLeuLeuProLysLeu 132  
 Db 329 TCGATTAACTAAACAGTCCCTATGGAAATGTACTTTCATGACCTCTCCATTGTC 388  
 QY 133 AlaThrSerGlyAspAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeu 152  
 Db 389 GTTGCTTCGGTGTATGATGGCAACTATGCACAACTGCGAGCTAGGACCTCTCATTTGT 448  
 QY 153 GlnSerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArgasp 172  
 Db 449 CAGGCAATTTCAAAGAGTTCATTATGAAAGTTGTAGCTGAAGTGAATTCAGGAA 508  
 QY 173 AlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeu 192  
 Db 509 TCTCCTCAAGACTATGAGCTTTAATTCGGTCTAAGGACAAAGAGGTTGATGAATTG 568  
 QY 193 LeuThrPheThrSerValGluGluThrValArgLys 204  
 Db 569 TTGACATTTAAGAGTGTGAANAGATGGTAAAGAA 604

## RESULT 7

BF009709 620 bp mRNA linear EST 06-DEC-2001  
 LOCUS ss83b10.y1 Gm-cl064 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-cl064-692 5' similar to TR:Q9S7H4 Q9S7H4 CHORISMATE MUTASE ;  
 mRNA sequence.  
 ACCESSION BF009709.1 GI:10709985  
 VERSION BF009709  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 620)  
 Shoemaker,R., Keim,P., Vodkin,L., Erpelnding,J., Coryell,V.,  
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N.,  
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,

TITLE  
JOURNAL  
COMMENT

McCann,R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cu@resgen.com  
 High quality sequence stop: 404.

## FEATURES

source  
 1..620  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /cultivar="Williams"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl064-692"  
 /tissue\_type="seedling epicotyls"  
 /dev\_stage="2 week old"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl064"  
 /notes="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 from the epicotyls of 2 week old seedling for the cultivar  
 Williams. The seedlings were germinated in a growth  
 chamber, excised above the soil level, and the plants  
 were placed in a 100 ppm solution of auxin for 24 hours  
 prior to harvesting. Complementary DNA was synthesized  
 from mRNA using a primer consisting of a poly(dT)  
 sequence with a XhoI restriction site. EcoRI adapters  
 were ligated to the blunt-ended cDNA fragments followed  
 by XhoI digestion. The cDNA fragments were directionally  
 cloned into the EcoRI-XhoI restriction site of the  
 pBluescript vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (GibcoBRL). This  
 library was constructed in the laboratory of Dr. Randy  
 Shoemaker."

## ORIGIN

Alignment Scores: 2,23e-83 Length: 620  
 Pred. No.: 735.00 Matches: 156  
 Score: 89.13% Conservative: 8  
 Percent Similarity: 84.78% Mismatches: 19  
 Best Local Similarity: 56.39% Indels: 2  
 Query Match: 10  
 DB: 10

US-10-624-061-16 (1-261) x BF009709 (1-620)

QY 1 MetAlaLysAlaAlaGluGlnSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20  
 Db 72 ATGCCCAAGCAGCAGACAAAGTCTGATCTGGAAATGTACACCGTAGCTTCAGTG 131  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleIleTyrGlyLeuIleGluArgAlaLys 40  
 Db 132 AGGGAGGATTTGGTTAGGCAAGAGGATACCATTTATGGTCTCATTCAGAGAGCCAA 191  
 QY 41 PheProSerAsnSerHisThrTyAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 Db 192 TTCCTTAGCAATCTCCACACCTATGATGAAGAGTAGTATGCTCAATCCAGGTTTGTGGC 251  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyArgTyr 80  
 Db 252 TCATTGGTAGAATTGTTCTTGAAGTACAGAGCCATTCAGCTAAGCTTGAAGATAC 311  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluGluAsnLeuProSerIleValProSer 100  
 Db 312 AAAAACCCCTGAAGAAACCCCTTCTTCCAGAAAAATTTACCAACCATCAATTGTGCCATCT 371

QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 DB 372 TACTCTCTCAACAGTTTTTGGATCTCTGAGCTGCTTCATTAACATTAAACAGTCCATC 431  
 QY 121 TrpLysMetTyrPheLysGlnLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
 DB 432 TGGAAAT-TATTTCCAGAGTTACTTTTCATTTGGTTGCTACTTCGGGGGATGAATCGAAC 490  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArg-ArgIleHi 160  
 DB 491 TATGCAACCAAGTACCTTATGATCTTTTCATTTATGGAGGCAATTTTATGGAGATTC 550  
 QY 160 sTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProIle 180  
 DB 551 CTTGGAAGAGTTTGTACCTGAGTGAGAAATCAAGGATTCCTCCAAACTAGGAAGCCTC 610  
 QY 180 uIleArgAla 183  
 DB 611 TATTCGAGCT 620

RESULT 8  
 BI420527 605 bp mRNA linear EST 15-AUG-2001  
 LOCUS LjN58758b9r Lotus japonicus node library 5 and 7 week-old Lotus  
 DEFINITION corniculatus var. japonicus cDNA 5', mRNA sequence.

ACCESSION BI420527 GI:15191550  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Lotus corniculatus var. japonicus (Lotus japonicus)  
 Lotus corniculatus var. japonicus  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;  
 Lotus.

REFERENCE 1 (bases 1 to 605)  
 Colebatch,G., Freund,S., Trevaskis,B and Udvardi,M.  
 Lotus japonicus root nodule ESTs: tools for functional genomics  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Udvardi MK  
 Molecular Plant Nutrition  
 Max Planck Institute of Molecular Plant Physiology  
 Am Muehlenberg 1, 14476 Golm, Germany  
 Fax: 49 331 567 8250  
 Email: udvardi@mpimp-golm.mpg.de

Seq primer: 17  
 High quality sequence stop: 605.  
 Location/Qualifiers  
 1..605  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="mRNA"  
 /cultivar="Gifu (B-129)"  
 /db\_xref="taxon:34305"  
 /dev\_stage="5 and 7 week-old plants"  
 /clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"  
 /note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI. The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

## FEATURES

## source

1..605  
 /organism="Lotus corniculatus var. japonicus"  
 /mol\_type="mRNA"  
 /cultivar="Gifu (B-129)"  
 /db\_xref="taxon:34305"  
 /dev\_stage="5 and 7 week-old plants"  
 /clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"  
 /note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI. The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,84e-79 Length: 605  
 Score: 723.50 Matches: 139  
 Percent Similarity: 86.74% Conservative: 18  
 Best Local Similarity: 76.80% Mismatches: 19  
 Query Match: 54.03% Indels: 5  
 DB: 12 Gaps: 1

US-10-624-061-16 (1-261) x BI420527 (1-605)

QY 1 MetAlaLysAlaAlaGluInSerProAspSerGlyAsnValTyrThrLeuAlaSerVal 20

DB 76 ATGCCCAAGCA-----GAATCAATGATGTGTTTCCCTGGAGTCTGTG 120  
 QY 21 ArgGluAspLeuValArgGlnGluAspThrIleLeuTyrGlyLeuIleGluArgAlaLys 40  
 DB 121 AGGAGAGATCTGTTAGCAAGAGATACCATGTTTTTGGTCTCATTTGACAGGGCTAGG 180  
 QY 41 PheProSerAsnSerHisThrTyrAspGluLysTyrAlaGlnIleGlnGlyPheCysGly 60  
 DB 181 TTCCCATGAATATATCACACCTATGATGATAAACTACTGGAAATCCCGGGTTTTTGTGGC 240  
 QY 61 SerLeuValGluPheValValLysAsnThrGluAlaIleGlnAlaLysAlaGlyValTyr 80  
 DB 241 TCATTTGGTCGAATGTTGTTTCAAAACACATGAGGCCATTCAAGCTATGCTGGAGGTAC 300  
 QY 81 LysAsnProGluGluAsnAlaPhePheProGluAsnLeuProProSerIleValProSer 100  
 DB 301 CAAACCCCTGAAGAAATGCTTCTTCCAGAACATTTACCATCACCATTGTGCCATCT 360  
 QY 101 TyrSerPheLysGlnPheLeuHisProGlyAlaAlaSerIleAsnIleAsnLysSerIle 120  
 DB 361 TACCCCTTCACACAGATTTTGCATCTCTGGAGCTGCTTCATTAACATAACAGTCGATT 420  
 QY 121 TrpLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
 DB 421 TGGAAATTAATTTGATGATGCTTCCCAATATTGTTGCTCGGTGATGATGCAAC 480  
 QY 141 TyrAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgAlaIleHis 160  
 DB 481 TATGCACAACTGCAGCTAGCGATCTTTCATTATTTGACGGCAATCTCTAGAGGATTCAT 540  
 QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
 DB 541 TATGGAAGTTTGTAGCTGAGCGGAATTCAGGAATCTCCTCAAGACTATGAGCCTTA 600  
 QY 181 Ile 181  
 DB 601 ATT 603

## RESULT 9

## AI973566

## LOCUS

## DEFINITION

AI973566 477 bp mRNA linear EST 30-NOV-2001  
 sc88e02.y1 Gm-cl019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
 Gm-cl019-435 5' similar to SW:CHMU\_ARATH P42738 CHORISMATE MUTASE  
 PRECURSOR ; mRNA sequence.

## ACCESSION

## VERSION

## SOURCE

## KEYWORDS

## ORGANISM

## Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

## REFERENCE

## AUTHORS

Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,  
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,  
 Beck,C., Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,  
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,  
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,  
 McCann,R., Waterston,R. and Wilson,R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 Contact: Shoemaker R/Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com

## TITLE

## JOURNAL

## COMMENT

Insert Length: 653 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 447.  
Location/Qualifiers

## FEATURES

source

1. .477  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-435"  
/tissue\_type="Immature seed coats of greenhouse grown plants"  
/lab\_host="DH10B (Gibco BRL)"  
/clone\_lib="Gm-cl019"  
/note="Vector: pSPORT1 (Life Technologies); Site 1: Not I; Site 2: Sal I; This cDNA library was constructed from mRNA isolated from immature seed coats (200-300 mgs) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly (dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells (Gibco BRL). This library was constructed by Dr. Lila Vodkin and Dr. Anu Khanna."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,628-79 Length: 477  
Score: 721.00 Matches: 141  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 53.85% Indels: 0  
DB: 9 Gaps: 0

US-10-624-061-16 (1-261) x At973566 (1-477)

QY 121 TPLysMetTyrPheLysGluLeuLeuProLeuLeuAlaThrSerGlyAspGlyAsn 140  
Db 1 TCGAAAGTATTTCAGAGAGTTACTTCATGCTTGTCTACTTCGGGTGATGGCAAC 60  
QY 141 TyrAlaGlnThrAlaAlaAspLeuSerLeuLeuGlnSerIleSerArgArgIleHis 160  
Db 61 TATGCGCAACTGCGAGCTTAATGACCTTCATTATTGCGAGTCCATCTCTAGAGGATTAC 120  
QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180  
Db 121 TATGGAAGTTTGTAGCTGAGTGGAATTCAGGATGCTCTCAAGACTAGAGCCCTTTA 180  
QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200  
Db 181 ATTCGAGCTAAGGATAAAGAGGATTGATGAATTTGTGCAATTCATCAAGCGTTGAAGAG 240  
QY 201 ThrValArgLysArgValGluLysIleAlaValValPheGlyGlnGluValAsnLeuAsn 220  
Db 241 ACGGTGAGAGAGAGTTGAAAGAGGCTGTGTGTGTGGCGAGGAGTGAATCTTAAC 300  
QY 221 SerAspAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240  
Db 301 AGTGATGACATGACATGAAGAAACCGTAATTTTCATCCATCAGTGGCTTCTAGCTGTAC 360  
QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLeuArgLeu 260  
Db 361 AAAAATGGGTGATACCTCTCACCAGAGGTTGAGTTGAGTACCTCTTGCGCCGCTCA 420  
QY 261 Asp 261  
Db 421 GAC 423

RESULT 10  
BE821647/c

## LOCUS

DEFINITION  
GN70001SAL0A11 Gm-r1070 Glycine max cDNA clone Gm-r1070-5781 3',  
mRNA sequence.  
ACCESSION  
BE821647  
VERSION  
BE821647.1 GI:10253881  
KEYWORDS  
EST.  
SOURCE  
Glycine max (soybean)

## REFERENCE

AUTHORS  
Vodkin,L., Keim,P., Shoemaker,R., Retzel,E., Khanna,A., Coryell,V.,  
Epeidig,J., Rapp,C., Shoop,E., Pardini,J., Liu,L. and Lewin,H.  
TITLE  
A Functional Genomics Program for Soybean (NSF 9872565)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Other ESTs: A1973566 corresponding to Gm-cl019-435 (5')  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vodkin@uiuc.edu

## FEATURES

source

1. .560  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="Gm-r1070-5781"  
/clone\_lib="Gm-r1070"  
/note="The library Gm-r1070 is a sequence-driven, reracked  
set of 9,216 clones selected from cDNA libraries from  
various tissues and stages of development of soybean that  
represent 2,639 sequences from immature cotyledons, 1,770  
from immature seed coats, 3,938 from flowers, and 869  
from young pods. The 5' ESTs of the source clones from  
the different libraries was used to select singletons, or  
a representative of each contig, which were reracked to  
form library Gm-r1070. The cDNA clones of the reracked  
Gm-r1070 library were then sequenced at the 3' end. The  
contig analysis to select unique genes was performed by  
the laboratory of Ernest Retzel, Center for Computational  
Genomics and Bioinformatics, University of Minnesota,  
http://www.cbc.umn.edu/researchprojects/soybean/index.html  
Reracking was performed by Genome Systems, St. Louis,  
http://www.genomesystems.com, and 3' sequencing by the  
Keck center for Comparative and Functional Genomics,  
University of Illinois,  
http://www.life.uiuc.edu/biotech/keck.html. Note: The  
corresponding 5' EST from each clone in the Gm-r1070  
library is listed in the 'OTHER EST' field. The detailed  
information on the source library for each clone can also  
be obtained by referring to the Genome Systems clone ID of  
the original cDNA library that is also listed under  
'OTHER EST'."

## ORIGIN

## Alignment Scores:

Pred. No.: 3,35e-79 Length: 560  
Score: 721.00 Matches: 141  
Percent Similarity: 100.00% Conservativity: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 53.85% Indels: 0

DB: 10 Gaps: 0

US-10-624-061-16 (1-261) x BE821647 (1-560)

QY 121 TPLysMetTyrPheLysGluLeuProLeuLeuAlaThrSerGlyAspGlyValAsn 140

DB 547 TGAATAATGTAATTCAGAGGTTACTTCCATTGCTTACTTGGGTGATGATGCAAC 488

QY 141 TTAAlaGlnThrAlaAlaAsnAspLeuSerLeuLeuGlnSerIleSerArgIleHis 160

DB 487 TATGCGCAAAATGACGCTAATGACCTTTCATTATTGACGCTCCATCTCTAGAAGGATTCA 428

QY 161 TyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTyrGluProLeu 180

DB 427 TATGGAAGTTTGTAGCTAGGTGAAATTCAGGATGCTCCCAAGACTACGAGCTTTA 368

QY 181 IleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSerValGluGlu 200

DB 367 ATTGAGCTAAGTAAGAAGATTGATGAAATTTGTCACATTTACAAGCGTTGAAGAG 308

QY 201 ThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluValAsnLeuAsn 220

DB 307 ACGGTGAGGAAGAGAGTTGAAAAGAGGCTGTGTGTTTGGGCGAGGAAGTCAATCTTA 248

QY 221 SerAspAsnAspAsnGluAsnArgLysPheAspProSerValAlaSerSerLeuTyr 240

DB 247 AGTGATGACAAATGACAAATGAAACCGTAATTTGATCCATGATGCTTCTAGCTTGTA 188

QY 241 LysAsnTrpValIleProLeuThrLysGluValGlnValGlnValGlnValGlnValGln 260

DB 187 AAAAAATGGGTGATGATCTCTACCAAGGAGGTTGAGTTGAGTACCTCTTGGCGCGCTTA 128

QY 261 Asp 261

DB 127 GAC 125

RESULT 11

CB827129

LOCUS 529 bp mRNA linear EST 16-MAY-2003

DEFINITION LjNEST7057r Lotus japonicus nodule library 5 and 7 week-old Lotus corniculatus var. japonicus cDNA 5', mRNA sequence.

ACCESSION CB827129

VERSION 1

KEYWORDS EST.

SOURCE Lotus

ORGANISM Lotus corniculatus var. japonicus (Lotus japonicus)

Lotus corniculatus var. japonicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.

REFERENCE 1 (bases 1 to 529)

Colbatch,G., Freund,S., Trevaskis,B and Udvardi,M.

Lotus japonicus root nodule ESTs: tools for functional genomics

Unpublished (2000)

Contact: Udvardi MK

Molecular Plant Nutrition

Max Planck Institute of Molecular Plant Physiology

Am Muehlenberg 1, 14476 Golm, Germany

Fax: 49 331 567 8250

Email: udvardi@mpimp-golm.mpg.de

Seq primer: T7

High quality sequence stop: 529.

Location/Qualifiers

1. 529

/organism="Lotus corniculatus var. japonicus"

/mol\_type="mRNA"

/cultiVar="Gifu (B-129)"

/db\_xref="taxon:34305"

/dev\_stage="5 and 7 week-old plants"

/clone\_lib="Lotus japonicus nodule library 5 and 7 week-old"

/note="Organ: Nodule; Vector: pSPORT1; Site 1: SalI; Site 2: NotI; The library was prepared using mRNA

FEATURES

source

extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."

## ORIGIN

Alignment Scores: 9,7e-79 Length: 529

Pred. No.: 717,00 Matches: 140

Score: 88.82% Conservative: 11

Best Local Similarity: 82.35% Mismatches: 19

Query Match: 53.55% Indels: 0

DB: 14 Gaps: 0

US-10-624-061-16 (1-261) x CB827129 (1-529)

QY 54 GlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGluAlaIle 73

DB 5 AAAATCCCGGGTTTTTGTGGCTCATTTGTCGAATGTGTGTTCAAAACACTGAGGCCATT 64

QY 74 GlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePheProGluAsnLeu 93

DB 65 CAAGCTATGCTGGAGGTACCAAAACCTGAAGAAATGCTTCTTCCAGACATTTA 124

QY 94 ProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAlaAlaSer 113

DB 125 CCATCAACCAATTGTGCCATCTTACCCCTTCACACAGATTTTGCATCTCTGAGCTGCTTC 184

QY 114 IleAsnIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuProLeuLeuAla 133

DB 185 ATTAACATAAAACAAATCGATTGGAATAAATTACTTTGATGAGTTGCTTCCAAATATTGTT 244

QY 134 ThrSerGlyAspAspGlyValAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeuGln 153

DB 245 CGTCCGGTGATGATGGCAACTATGCAAACTGCAGTAGCGATCTTTCATTATTGCAG 304

QY 154 SerIleSerArgArgIleHisTyrGlyLysPheValAlaGluValLysPheArgAspAla 173

DB 305 GCAATCTCTAGAAGGATTCAATTATGGAAGTTTGTAGCTGAGGCGAAATTCAGGAACTCT 364

QY 174 ProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeu 193

DB 365 CTTCAAGACTATGAGCTTTTAAATTCGTCTAAGCACACAGAGGTTGATGAATATTG 424

QY 194 ThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaValValPhe 213

DB 425 ACATTTGAGAGTGTTCAGAGATGTTGATAAAGAGAGTTGAAAAGAGGCCATGTTGTT 484

QY 214 GlyGlnGluValAsnLeuAsnSerAspAsp 223

DB 485 GNGCAGGAGGTAGGCTTAACAGTGATGAT 514

## RESULT 12

CB8239446

LOCUS

DEFINITION

RG239446 422 bp mRNA linear EST 28-NOV-2001

sab70d03.y1 Unknown Library Type Glycine max cDNA clone GENOME

SYSTEMS CLONE ID: Gm-c1032-2453 5' similar to TR:Q9S7H4 Q9S7H4

CHORISMATE MUTASE ; mRNA sequence.

ACCESSION

RG239446

VERSION

RG239446.1

KEYWORDS

SOURCE

ORGANISM

Glycine max (soybean)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 422)

Shoemaker,R., Keim,P., Vodkin,L., Expelding,J., Corvelli,V.,

Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,

Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,

Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,

Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,

McCann,R., Waterston,R. and Wilson,R.



SAIL\_30\_H03\_g1\_A002 Salicylic acid-treated seedlings Sorghum bicolor  
 cDNA clone SAIL\_30\_H03\_A002 5', mRNA sequence.  
 CD427426  
 CD247426.1 GI:31333689  
 EST.  
 Sorghum bicolor (sorghum)  
 Sorghum bicolor  
 Sorghum bicolor  
 Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD  
 clade; Panicoideae; Andropogoneae; Sorghum.  
 1 (bases 1 to 810)  
 Cordonnier-Pratt, M.-M., Wentzel, V., Suzuki, Y., Sugano, S.,  
 Klein, R.R., Liang, C., Sun, F., Sullivan, R., Shah, M., Salzman, R.,  
 Chua, Tan, N., Gonzalez, M., Lane, S., Miller, V., Nanda, P.,  
 Olasinde, O., Eastman, A. and Pratt, L.H.  
 An EST database from Sorghum: salicylic acid-treated seedlings  
 Unpublished (2003)  
 Other\_ESTs: SAIL\_30\_H03.bl\_A002  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mprattd@uga.edu  
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in  
 the Human Genome Center, University of Tokyo Institute of Medical  
 Science; plant material and RNA prepared at Texas A & M University;  
 sequencing done in the Laboratory for Genomics and Bioinformatics,  
 University of Georgia. Sequence ends have been trimmed to exclude  
 vector and regions below Phred quality 16. Three-prime sequences  
 are presented as their reverse complement and have been trimmed to  
 exclude polyA.  
 Seq primer: Sug5 (CTTCTGCTCTAAAGTCGG).

```

seq primer: 50ys (CATTCTTCTATTAAGACGTCGGT).
Location/Qualifiers
    1..810
        /organism="Sorghum bicolor"
        /mol_type="mRNA"
        /cultivar="BTx623"
        /db_xref="taxon:4558"
        /clone="SA_30_H03_A002"
        /lab_host="DH10B-T1 phage-resistant E. coli"
        /clone_lib="Salicylic acid-treated seedlings"
        /note="Vector: pME18S-F13; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from seedlings grown in hydroponic culture. At 8 days of age, medium was supplemented with 1 mM salicylic acid (SA). Roots and shoots were harvested after 27 and 72 hr and material from both time points was combined prior to RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraII sites of the pME18S-F13 vector (5-prime DraII site is CACTGTGTG, 3-prime DraII site is CACCATGTG). XhoI excises the cDNA insert."
ORIGIN
Alignment Scores:
Pred. No.:      8,92e-74      Length:      810
Score:          679.50       Matches:     139
Percent Similarity: 69.84%   Conservative: 37
Best Local Similarity: 55.16% Mismatched: 71
Query Match:     50.75%   Indels:      5
                  14      Gaps:         3
US-10-624-061-16 (1-261) x CD427426 (1-810)
12 GlyasnvAlTyThrLeuAlaserValArgGluspsLeuValArgcIngIuasptRlle 31
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
26 GGCAGCACTAGAGCTGCGCGGTGCCGATCCGCTGCTCGCGTGAGGACTCCGTG 85
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
32 IleTgLyLeuIcLuarGalalyPheProSerAsnSerHisThrTyAspGluLlys 51
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
86 GTCTTCGCTCATCGAGCGGCGCCGCGATCCGCGGAACCGCGCGGCTACGCGCCGCC 145
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```



QY 52 TyrAlaGlnIleGlnGlyPheCysGlySerLeuValGluPheValValLysAsnThrGlu 71  
 Db 146 GCGCGGCGGCGGAGGATAT-----TCGCTCGTGGAGTTCTTCTCGGGAAGCAGAG 199  
 QY 72 AlaIleGlnAlaLysAlaGlyArgTyrLysAsnProGluGluAsnAlaPhePheProGlu 91  
 Db 200 GCGCTCAAGCGGAGGCTGCATATCAAAAGCCAGAGATGTTCCATCTTCCTCAA 259  
 QY 92 AsnLeuProProSerIleValProSerTyrSerPheLysGlnPheLeuHisProGlyAla 111  
 Db 260 GATCTTCCTCGGCTCTCTTCTACCAAGCCTTCCCAAGGCTTTGCACCTTTTGT 319  
 QY 112 AlaSerIleAsnLysSerIleTyrLysMetTyrPheLysGluLeuLeuProLeu 131  
 Db 320 TTGTTGGTCACTGTAATGCAATATGAAATATATTTTGTGAATGTCTCCATTA 379  
 QY 132 LeuAlaThrSerGlyAspGlyAsnTyrAlaGlnThrAlaAlaAsnAspLeuSerLeu 151  
 Db 380 TTCACCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 439  
 QY 152 LeuGlnSerIleSerArgIleHisTyrGlyPheValAlaGluValLysPheArg 171  
 Db 440 CTCGAGTTCTATCAAAAGATTCATATTTGGTAAATATGTTGCTGAGGTCAAGTTCAA 499  
 QY 172 AspAlaProGlnAspTyrGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLys 191  
 Db 500 GATGCTCTCAAGATTATGCTGATTAATAAGCAAGACAGCAATCTTTGATGGAT 559  
 QY 192 LeuLeuThrPheThrSerValGluGluThrValArgLysArgValGluLysLysAlaVal 211  
 Db 560 CTGCTCACATTCAAGCTGTGGAAGAGAGAGTTTAAAGAGAGATGACAGAAAGCCAGG 619  
 QY 212 ValPheGlyGlnGluValAsnLeu-----AsnSerAspAspAsnAspAsnGluAsnArg 229  
 Db 620 ACCTTCGGCGAGAACCTCACTTCGAGGACAATGCCACTGCCAGTACAGCGAA---TGC 676  
 QY 230 LysPheAspProSerValAlaSerSerLeuTyrLysAsnTyrValIleProLeuThrLys 249  
 Db 677 AAGTCAATCTTAAGTGTCTCCAGCTGATGATGATGATGATGATGATGATGATGAT 736  
 QY 250 GluValGlnValGluTyrLeuLeuArgArgLeuAsp 261  
 Db 737 GAGTCGAGTCGAGTATCTCTCGCGCGGCTCGAC 772  
 RESULT 14  
 LOCUS BM093460  
 DEFINITION saJ09c02.y1 Gm-c1065 Glycine max cDNA clone GENOME SYSTEMS CLONE  
 ID: Gm-c1065-9724 5', similar to TR:Q957H4 Q957H4 CHORISMATE MUTASE  
 ;, mRNA sequence.  
 ACCESSION BM093460  
 VERSION BM093460.1 GI:17022426  
 KEYWORDS EST.  
 SOURCE Glycine max (soybean)  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 1 (bases 1 to 560)  
 REFERENCE  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McCann, R., Waterston, R. and Wilson, R.  
 PUBLIC SOYBEAN EST PROJECT  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800

Fax: 314 286 1810  
 Email: eat@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: ccu@resgen.com  
 High quality sequence stop: 421.  
 Location/Qualifiers  
 1..560  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1065-9724"  
 /tissue\_types="germinating shoots"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-c1065"  
 /note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2:  
 XhoI; The cDNA library was constructed from mRNA isolated  
 germinating shoots of the cultivar Williams. The seeds  
 were allowed to germinate for 24 hours prior to being  
 cold stressed for 2 days at 4C. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with a XhoI restriction site. EcoRI  
 adapters were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA fragments were  
 directionally cloned into the EcoRI-XhoI restriction site  
 of the pBluescript vector. The ligated cDNA fragments were  
 transformed into DH10B host cells (GibcoBRL). This library  
 was constructed in the laboratory of Dr. Randy  
 Shoemaker."

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,38e-73 Length: 560  
 Score: 672.00 Match: 140  
 Percent Similarity: 84.85% Conservative: 0  
 Best Local Similarity: 84.85% Mismatches: 0  
 Query Match: 50.19% Indels: 25  
 DB: 12 Gaps: 1  
 US-10-624-061-16 (1-261) x BM093460 (1-560)  
 QY 122 LysMetTyrPheLysGluLeuProLeuLeuAlaThrSerGlyAspAspGlyAsnTyr 141  
 Db 3 AAAATGATATTTCAAAGAGTTTACTTCCATTGCTTCTCGGTTGATGATGGCACTAT 62  
 QY 142 AlaGlnThrAlaAlaAsnAspLeuSerLeu----- 151  
 Db 63 GCGCAACTGCGCTAATGACCTTTTCATTATTCGAGGACCAACTCAAGATTTCGTAGTC 122  
 QY 152 -----LeuGlnSerIleSerAr 157  
 Db 123 TCCTAAACTAGTATTGATCTCTGACTGATCTTATATAAACAATTTCAGTCCATCTCTAG 182  
 QY 157 GAGTLeHisTyrGlyLysPheValAlaGluValLysPheArgAspAlaProGlnAspTy 177  
 Db 183 AAGGATTCACTATGAAAGTTTGTAGCTGAGTGAATTCAGGATGCTCTCTCAAGACTA 242  
 QY 177 rGluProLeuIleArgAlaLysAspLysGluGlyLeuMetLysLeuLeuThrPheThrSe 197  
 Db 243 CGAGCTTTAATTCGAGCTAAGGATAAAGAGGATTGATGAATTTGTTGACATTACAAAG 302  
 QY 197 rValGluGluThrValArgLysArgValGluLysLysAlaValValPheGlyGlnGluVa 217  
 Db 303 CGTTGAAGAGACGGTGAAGAGAGAGAGTGTGAAAGAAAGGCTGTGTGTTTGGCGAGGAAGT 362  
 QY 217 LasnLeuAsnSerAspAspAsnAsnGluAsnArgLysPheAspProSerValAlaLase 237  
 Db 363 GATCTTTAACAGTGTGATGACATGACATGAATGAACCGTAAATTTGATCCATCAGTGGCTTC 422  
 QY 237 rSerLeuTyrLysAsnTrpValIleProLeuThrLysGluValGlnValGluTyrLeuLe 257  
 Db 423 TAGCTTTGTACAAAAATTTGGTGATACCTCTCACCAGAGGTTTCAGGTTGAGTACCTCTT 482

